தமிழுக்கான எழுத்துப்பிழைத் திருத்தியும் இலக்கணப்பிழைத் திருத்தியும் உருவாகுவதற்கான நுணுக்கங்கள்

(Nuances of Making Spell and Grammar Checker for Tamil)

Prof. Rajendran Sankaravelayuthan

Amrita University Coimpatore

rajushush@gmail.com

&

Dr. R. Amutha

amutham.nilavu@gmail.com

Senior Resource Person, CIIL, Mysore

ABSTRACT

The monograph "Nuances of making Spell and Grammar Checker for Tamil" written in Tamil aims at describing the nuances of making a spell and grammar checker for Tamil. The book has 5 chapters.

Chapter 1: The first chapter is introduction. It introduces the topic of the research work. It explains the aims and objectives of the book. It also discusses briefly about the earlier works on this topic and also about the use of such research work.

Chapter 2: The second chapter discusses about the nuances of spell and grammar checking.The possibility of occurrence of spelling errors, finding out the spelling errors and correcting them, finding out the non-words in language and correcting them and the methods of correcting spelling mistakes using Bayesian method and short distance method are discusse under the subheading 'spell checking'. Under the subheading 'grammar checking' the following points have been discussed: syntax based checking, statistical based checking and rule based checking. The background principles, part of speech tagging, chunking, grammar errors, non-specific modifier and the word followed, question, other mistakes, finding out the boundary of sentences, controlled language

correction and style correction have been discussed under the subheading 'Style and grammar checker'.

Chapter 3: This chapter discusses about steps or initiatives to be taken for spelling checking and grammar checking. A description is made about the phonological structure of Tamil including vowel and consonant phonemes and their phonotactic structure and consonant clusters and syllabic structure. This is followed by morphological structure of Tamil. Under morphological structure of nouns, the inflection in nouns, noun stems, inflected stems which includes inflected stem ending in ttu, aRRu, stems formed by doubling of consonants, stems of pronouns, number marking, augmentation by in and an, case suffixes and sandhi rules have been discussed. Verb inflection, verb stems, verb roots, causative and non-causative verb stems, derived verb stems, soft verb stems, middle verb stems and hard verb stems, finite verb stems, imperative verb stems, interrogative verb stems, tense suffixes, person-number-gender suffixes, inflectional paradigms, irregular verbs, optative verb forms, non-finite verb forms, infinitive verb forms, adverbial and conditional verb forms, adjectival verb forms, nominalized verbs, verbal nouns and appellative verbs are discussed under the morphological structure of verbs. This is followed by a section on computational morphology. Under this heading theory and implementation, formalism and construction, frame work for hybrid model, allomorphic state, nouns, pronouns, verbs, adjectives, implementation of hybrid model. are elaborated on. Next is on sandhi rules. This is followed by spell correction in Tamil. Under this headings mistakes due to pronunciations, mistakes due to not knowing phonological structure, mistakes due to interference of spoken Tamil, mistakes due to mother tongue influence and morphological mistakes are discussed.

Chapter 4: The fourth chapter is on making spell and checker for Tamil. Under this heading the spell checker made for Tamil have been discussed first. The spell checkers prepared by Danabalam, Vijay Sankar Ram and Deivasundram have been discussed.

Chapter 5: Conclusion,

தமிழுக்கான எழுத்துப்பிழைத் திருத்தியும் இலக்கணப்பிழைத் திருத்தியும்

(Spell and Grammar Checker for Tamil)

ஆக்கியோர்

பேராசிரியர் இராசேந்திரன் சங்கரவேலாயுதன்

அமிர்தா விஷ்வவித்யபீடம் பல்கலைக்கழகம், கோயம்புத்தூர்

rajushush@gmail.com

முனைவர் இரா. அமுதா

முதுநிலை மூலவள ஆய்வுஞர், இந்தியமொழிகளின் மையநிறுவனம், மைசூர்

amutham.nilavu@gmail.com

கோயம்புத்தூர்

பெப்ருவரி 2020



என்னுரை

தமிழில் எழுத்துப்பிழைத் திருத்தியும் இலக்கணப்பிழைத் திருத்தியும் ((Spell and Grammar Checker for Tamil) என்ற நூல் பல்கலைக்கழக நிதிநல்கையில் 1997-2001 என்ற காலகட்டத்தில் நானும் பேராசிரிய எஸ். பூபதி அவர்களும் இணைந்து செய்த ஆய்வின் அடிப்படையில் எழுதப்பட்ட நூலாகும். இத்திட்டத்தில் தற்போது ஹைதராபாத் மையப் பல்கலைக்கழகத்தில் இணைப்பேராசிரியராகப் பணியாற்றுகின்ற முனைவர் எஸ். அருள்மொழி அவர்கள் ஆய்வாளராகப் பணியாற்றினார். ஹைதராபாத் மையப் பல்கலைக்கழகத்தில் பேராசிரியராகப் பணியாற்றும் கவி நாராயணமூர்த்தி அவர்கள் ப்ரோலாக் மொழியில் இதற்கான நிரலை எழுதி உதவினார். முனைவர் ஆர். அமுதா பின்னர் இத்திட்டத்தில் இணைந்துகொண்டார். கடந்த சில ஆண்டுகளாக எனது மடி(க்கணி)யில் உறங்கிக்கொண்டிருந்த இவ்வாய்வேட்டை Language in India என்ற தமது மின் திங்கள் இதழில் நூலாக வெளியிட பேராசிரியர் திருமலை அவர்கள் உதவிக்கரம் நீட்டியதால் பல மாற்றங்களை மேற்கொண்டு இந்நூல் வெளியிடப்படுகின்றது.

அன்புடன்

ச.இராசேந்திரன்

பொருடக்கம்

Key words:spell checker, syntax based checking, rule based checking, parts of speech tagging, phrase chunking, syllable structure, phonotactics, morphotactics, documents, OCR, STT, VRS, typographicl errors, pronunciation variation, non-word error detection, isolated-word error correction, context-dependent error detection and correction, typographic errors, cognitive errors, deletion error, insertion error, substitution error, transposition error, Beyesian rules, noisy channel model, probability theory, voice recognition, minimum edit distance, statistics-based checking, morphological analyzer, two level morphology, paradigmatic approach, Levenstein distance, corpus driven approach, machine learning approach, hybrid approach,

இயல் 1

அறிமுகம்

தமிழுக்கு ஒரு எழுத்து மற்றும் இலக்கணப் பிழை திருத்தி உருவாக்குவதற்கான சாத்தியங்களை ஆய்வதும் அதற்கான முன்னேற்பாடுகளைச் செய்வதும் தான் இந்நூலின் நோக்கம் ஆகும். மேலும் எழுத்து மற்றும் இலக்கணப் பிழை திருத்தி தொடர்பான தகவல்களை எல்லாம் திரட்டித் தமிழில் தருவது இதன் துணை நோக்கமாகும். தமிழுக்கான ஒரு எழுதுப் பிழை மற்றும் இலக்கணப் பிழை திருத்தும் ஒரு முமையான மென்பொருள் உருவாக இங்கு முயற்சிக்கப்படவில்லை. இதற்கான முன்னேற்பாடுகளையும் கோட்பாடுகளையும் மாதிரியையும் முன்வைப்பது தாண் இதன் நோக்கம்.

தமிழ் போன்ற ஒட்டுநிலை மொழிகளின் உருபனியல் அமைப்பு கலவைத்தன்மை உடையதாகும். இத்தகைய மொழிகளில் எழுத்துப் பிழைதிருத்தி உருவாக்குவது சிக்கலான செயல்பாடாகும். தமிழுக்கென்று தனித்தன்மையான ஒலியன் அமைப்பொழுங்கும் (Phonotactics) உருபன் அமைப்பொழுங்கும் (Morphotactics) உள்ளன. எழுத்துப்பிழை திருத்தியும் இலக்கணப்பிழை திருத்தியும் இவ்வமைப்பு ஒழுங்குகளை உள்வாங்கிச் செயலாற்ற வேண்டும். தமிழ் வாக்கியங்களில் பயனிலையில் வரும் வினைகள் அதன் விகுதி (இடம்-பால்-எண் விகுதி) வழி வாக்கியத்தில் உள்ள எழுவாயுடன் உடன்பட்டு வரும். மேலும் வினையின் பங்கெடுப்பாளர்களாக வரும் பெயர்ச்சொற்கள் அவற்றின் பங்களிப்புகளுக்கு ஏற்றவாறு வேறுபட்ட வேற்றுமை உருபுகளை ஏற்றுவரும். இத்தகைய இலக்கணக்கூறுகளை உள்ளடக்கிய இலக்கணப்பிழை ஆய்வியை உருவாக்குவது ஒரு அறைகூவலான செயல்பாடாகும். தமிழ் ஒரு வினை இறுதி மொழியாகும்; அதாவது SOV வரிசை அமைப்புள்ள மொழியாகும். எழுவாயும் செயப்படுபொருளும் முன்பின்னாக மாறிவர இயலும் (SOV, OSV). இருப்பினும் எழுவாயாகவும் செயப்படுபொருளாகவும் செயல்படும் பெயர்த் தொடர்களுக்குள் வரும் அடைகொளி அடை, பெயரடை, எண்ணடை, அளவையடை போன்ற இலக்கணக்கூறுகள் ஒரு குறிப்பிட்ட வரிசைமுறையிலேயே வரும். இத்தொடரியல் அமைப்பையும் இலக்கணத்திருத்தி உள்வாங்கிச் செயலாற்ற வேண்டும்.

எழுத்துப்பிழை திருத்தியும் இலக்கணப்பிழை திருத்தியும் உருவாக்குவதற்கான நடைமுறைகளைப் பயன்படுத்தி பரிசோதனை நிலையில் தமிழுக்கு ஒரு எழுத்துப்பிழை திருத்தியையும் இலக்கணத் திருத்தியையும் மாதிரியாக உருவாக்க முயற்சி எடுக்கப்பட்டுள்ளது. புள்ளியியல் நெறிமுறைகளும் இவ்வாய்வேட்டில் பயன்படுத்தப்பட்டுள்ளது. தவறான சொல் சரியான சொல்லிருந்து எவ்வளவு விலகி இருக்கின்றது, திருத்து தூரம் எவ்வளவு என்ற தகவல்கள் புள்ளியியல் நெறிமுறை அடிப்படையில் பெறப்பட்டுள்ளன. தமிழில் பிழை ஆய்வுகளுக்கான தரவுகள் சேகரிக்கப்பட்டு ஆயப்பட்டு பிழைகளின் வகைகளும் நிகழ்வெண்களும் கண்டுபிடிக்கப்பட்டுள்ள. இத்தரவுகள் மாணவர்களின் தேர்வு ஆவணங்களிலிருந்து கட்டுரைகளிலிருந்தும் பிற மூலங்களிலிருந்தும் சேகரிக்கப்பட்டுள்ளது.

முன்னோடிகள்

தனபாலன், ரஞ்சனி பார்த்தசாரதி, கீதா என்போர் தமிழ் எழுத்துப்பிழை திருத்தி ஒன்றை உருவாக்கியுள்ளனர். இதுகுறித்து நான்காவது இயலில் கூறப்பட்டுள்ளது. இது போன்று விஜய் சுந்தர் ராம் அவர்களும் தமிழுக்கு என்று ஒரு எழுத்துப்பிழை திருத்தி உருவாக்கியுள்ளார்கள். இதன் விரிவான விளக்கமும் நான்காவது இயலில் விளக்கப்பட்டுள்ளது. தமிழ் சொல்லாய்விகளை உருவாக்கியுள்ளவர்களும் தங்களது மென் பொருள்களில் எழுத்துப் பிழை திருத்தத்திற்கான வசதி செய்துள்ளனர். இதில் தெய்வ சுந்தரம் அவர்கள் உருவாக்கி தமிழ்ப் பேராயம் வெளியிட்டுள்ள “மென்தமிழ் தமிழ்ச் சொல்லாளர்” மிகவும் முக்கியத்துவம் வாய்ந்தது. இது சொற்பிழை திருத்தி, சந்திப்பிழை திருத்தி என்ற வசதிகளைக் கொண்டுள்ளது. இது குறித்தும் ஐந்தாவது இயலில் குறிப்பிடப்பட்டுள்ளது.

பயன்பாடு

தமிழுக்கு எழுத்துப்பிழை திருத்தியும் இலக்கணப் பிழை திருத்தியும் மாதிரியாக வடிவமைக்கப்பட்டுள்ளது. இதை விரிவாக்கி தமிழுக்கு ஒரு முழுமையான எழுத்து மற்றும் இலக்கணப்பிழை திருத்தி உருவாக்க இயலும்.

இவ்வேடு தமிழில் ஒரு நல்ல எழுத்துப்பிழை திருத்தியும் இலக்கணப்பிழை திருத்தியும் உருவாக்குவதற்கு முன்னோடியாக அமையும். தமிழில் எழுத்துப்பிழை திருத்தம் பற்றியும் இலக்கணப்பிழை பற்றி பேசும் நூல்களும் கட்டுரைகளும் பல இருந்தாலும் இவற்றைப் பயன்படுத்தி எவ்வாறு ஒரு எழுத்துப் பிழை திருத்தி உருவாக்குவது மற்றும் இலக்கணப் பிழை திருத்தி உருவாக்குவது என்பது குறித்த கட்டுரைகளும் நூல்களும் மிகக் குறைவு; ஏன் இல்லை என்றே சொல்லலாம். இநூல் இந்த இடைவெளியை நிரப்பும்.

இயல் 2

எழுத்துப்பிழை திருத்தமும் இலக்கணப்பிழை திருத்தமும்

2.0 முன்னுரை

சொல்லாய்வியின் வழி ஆவணங்கள் (Documents) உருவாக்கப் படுகின்றன. ஆவணம் ஒரு கடிதமாகவோ சிறுகுறிப்பாகவோ கட்டுரையாகவோ நூலாகவோ இருக்கலாம். சொல்லாய்வியில் ஆவணத்தை உருவாக்கப் பொதுவாகப் பின்வரும் மூன்று முறைகள் பயன்படுத்தப்படுகின்றன:

1. கணிப்பொறியின் விசைப்பலகை (Keyboard) வழி உள்ளீடு செய்தல்
2. பேச்சுணரி ஒழுங்கு முறையைப் (Voice Recognition System (VRS)) பயன்படுத்தி உரையை உள்ளீடு செய்தல்: இவ்வாறு செய்யும் பொழுது ஒலிவடிவில் பேசப்படும் செய்திகள் உரைவடிவிற்கு மாற்றப்படுகின்றன. இதற்குப் பேச்சிலிருந்து உரை (Speech to Text (STT)) என்ற தொழில் நுட்பம் உதவுகிறது.
3. தட்டச்சு செய்தல் அல்லது அச்சடிக்கப்பட்ட உரையை ஒளிவழி எழுத்துணர்தல் (Optical Character Recognition (OCR)) என்ற தொழில்நுட்பத்தின் மூலம் களிப்பொறிக்கு உள்ளீடு செய்து சொல்லாய்விற்கு ஆவணமாக மாற்றுவது.

உரையை உள்ளீடு செய்ய மேற்கண்ட மூன்று முறைகளுள் ஏதேனும் ஒன்றைப் பின்பற்றலாம். பின்னர் சொல்லாய்வியின் மூலம் அவைகளைக் கையாளலாம். உரைக்குள் எழுத்துப் பிழைகளும் இலக்கணப் பிழைகளும் இருப்பதற்கான வாய்ப்புகள் உள்ளன. ஆங்கில உரைகளில் உள்ள எழுத்துப் பிழைகளைக் கணிப்பொறிதானாகவே புரிந்துக்கொண்டு மாற்றுவதற்குச் சொல்லாய்வியில் வசதிகள் உள்ளன. இதற்கு எழுத்துப் பிழைகளைக் கண்டறிந்து திருத்துதல் என்று பெயர். இத்தகைய பிழைதிருத்தி தமிழ் சொல்லாய்விகளில் உருவாக்க முயற்சிகள் எடுக்கப்பட்டு வருகின்றன. எழுத்துப் பிழைகள் நிகழ்வதற்கான வாய்ப்புகள், எழுத்துப் பிழைகளின் வகைகள், பிழைகளைக் கண்டறிந்து திருத்துவதற்கான செயல்பாட்டு வழிமுறைகள் போன்றவையெல்லாம் விரிவாக ஆராய்ந்து செயல்படவேண்டி உள்ளது.

பிழைகளை எழுத்துப்பிழைகள், இலக்கணப்பிழைகள் என்று வரையறுப்பது கவனக்குவிப்பு செய்வதற்கு எளிமையாக அமைந்தாலும் இரண்டையும் வெவேறான

செயல்பாடுகள் என்று கூறுவது கடினம். எழுத்துப்பிழைகள் இலக்கணப்பிழைகளுக்கு வழிவகுக்கலாம் அல்லது இலக்கணப்பிழைகள் எழுத்துப்பிழைகளாகத் தோன்றலாம். குறிப்பாகத் தமிழ் போன்ற திரிபு ஒட்டுகள் மிகுந்த மொழியில் எழுத்துப்பிழை இலக்கணப்பிழைக்கும் இலக்கணப்பிழை எழுத்துப்பிழைக்கும் வழிவகுக்கும். சந்திப் பிழைகளை அல்லது புணர்ச்சிப் பிழைகளை எழுத்துப்பிழைகளாகவும் இலக்கணப்பிழைகளாகவும் கருதலாம்.

2.1. எழுத்துப்பிழை திருத்தம்

எழுத்துப்பிழை திருத்தம் எழுத்துப்பிழை நிகழ்வதற்கான வாய்ப்புகளைக் கண்டறிதல், எழுத்துப்பிழைகளக் கண்டறிதல், பிழை நீக்கும் வழிமுறைகளைக் கண்டறிதல், பிழை நீக்குவதில் உள்ள சிக்கல்களை அறிதல் என்பன போன்ற செயல்பாடுகளை உட்படுத்தும்.

2.1.1 எழுத்துபிழைகள் நிகழ்வதற்கான வாய்ப்புகள்

உரைச் செய்திகளைக் கீழ்கண்ட ஏதேனும் ஒருங்குமுறைகளில் உள்ளீடு செய்யப்படும்போது எழுத்துப்பிழைகள் நிகழ்வதற்கு வாய்ப்புகள் உள்ளன.

- விசைப்பலகை வழி உள்ளீடு செய்யும்பொழுது தட்டச்சுப் பிழைகள் (Typographical Errors) ஏற்படுகின்றன.
- பேச்சுணரி ஒழுங்குமுறை மூலம் உள்ளீடு செய்கின்றபொழுது சொல்லின் உச்சரிப்பு மாற்றங்களினால் பிழைகள் ஏற்படுகின்றன. இத்தகைய பிழைகளுக்கு உச்சரிப்பு வேறுபாடு (Pronunciation Variation) என்று பெயர்.
- ஒளி எழுத்துணரி (OCR) வழி உள்ளீடு செய்கின்றபொழுது கருவிகளின் குறைபாடுகளினாலோ அல்லது எழுத்துக்களைச் சரியாகக் கருவிகள் படித்தறிய இயலாததாலோ பிழைகள் தோன்றுகின்றன. எடுத்துக்காட்டாக ஆங்கிலத்தில் I என்ற எழுத்திற்கும் 1 என்ற எழுத்திற்கும் வேறுபாடு தெரியாததால் பிழை ஏற்படுகின்றன. பின்வரும் அட்டவணையில் சில சரியான எழுத்துக்களும் அவற்றிற்கிணையான தவறாகப் புரிந்துகொள்ளப்படும் எழுத்துக்களும் பட்டியலிடப்பட்டுள்ளன:

சரியான எழுத்து	தவறாகப் புரிந்துகொள்ளப்பட்ட எழுத்து
I	1
D	0
ri	n
0	θ

மேற்கண்டவாறு மூன்றுவகையான உள்ளீட்டு முறைகளிலும் அதற்கான கருவிகளைப் பயன்படுத்தும் பொழுது பிழைகள் தோன்றுகின்றன. பேச்சறிதல் ஒழுங்குமுறை மூலம் உள்ளீடு செய்கின்றபொழுது ஒரே வார்த்தையை அதன் இலக்கணக் கூறு மற்றும் தன்மைகளுக்கு ஏற்ப வெவ்வேறு விதமாக உச்சரிக்க வேண்டியுள்ளது. எடுத்துக்காட்டாக ஆங்கிலத்தில் Object என்ற சொல் பெயர்ச் சொல்லாகவும் வினைச் சொல்லாகவும் பயன்படுகின்றது. பெயர்ச்சொல்லாக பயன்படும்போது 'ஆப்ஜிக்ட்' என்றும் வினைச் சொல்லாகப் பயன்படுத்தும்போது 'அப்ஜெக்ட்' என்றும் உச்சரிப்பு செய்யவேண்டியுள்ளது. அதைப்போல் The என்ற சொல்லைச் சூழலுக்கு ஏற்ப [ðə] என்றும் என்றும் [ði] மாறுப்பட்ட இருவகைகளில் உச்சரிக்க வேண்டியுள்ளது. இவ்வாறாக, ஒரே சொல்லிற்கான உச்சரிப்பு மாற்றங்களுக்குத் திட்டவட்டமாக நன்கு வரையறுக்கப்பட்ட விதிமுறைகள் உள்ளன. அவ்விதி முறைகளைக் கணிப்பொறிக்கு ஏற்ற வண்ணம் வடிவமைப்புகளாக வடிவமைத்து இத்தகைய பிழைகள் நீக்கப்படுகின்றன.

2.1.2 எழுத்துப் பிழைகளைக் கண்டறிதல்- சரிசெய்தல் சிக்கல்களின் வகைகள்

குகிச் (Kukich, 1992) தமது மதிப்பீட்டாய்வு கட்டுரையில் எழுத்துப்பிழை திருத்தத்தை பரந்த அளவில் மூன்றாகப் பிரிகின்றார். இதன் படி எழுத்துப்பிழைகளைக் கண்டறிதல்-சரிசெய்தல் தொடர்பான சிக்கல்கள் பின்வரும் செயல்பாடுகளை உள்ளடக்கும்:

- மொழியில் இல்லாத சொற்களைக் கண்டறிதல் (Non-Word Error Detection)
- எழுத்துப்பிழை உள்ள செய்திகளை மட்டும் தனிமைப்படுத்தி ஆராய்தல் (Isolated-word error correction)

- எழுத்துப்பிழை உள்ள சொல்லின் முன்பின் உள்ள அல்லது அச்சொல் இடம்பெற்ற தொடரின் பொருளையமாகக் கொண்டு பிழைநீக்கி சரிசெய்தல் (context-dependent error detection and correction)

குகிச் (Kukich 1992) மனிதர்களின் தட்டச்சுப் பிழைகளை இரு பிரிவுகளாகப் பிரிக்கிறார்:

1. தட்டச்சுப் பிழைகள் (typographic errors)
2. புரிதல் பிழைகள் (cognitive errors) சொல்லின் எழுத்து வடிவம் உரையாளருக்குத் தெரியாததால் ஏற்படும் பிழைகள்

டமரு (Dameraru, 1964) தவறாக எழுதப்பட்ட சொற்களில் 80% தனிநிலைத் தவறால் ஏற்பட்ட பிழைககளாகக் காண்கிறார். இதன் படி தட்டச்சுப் பிழைகளைப் பின்வரும் நான்கு வகைகளாகப் பிரிக்கலாம்:

1. விடுபடு பிழை (Deletion Error): இப்பிழையில் ஒரு எழுத்து விடுபட்டிருக்கும். எடுத்துக்காட்டாக, *சட்டம்* என்ற சொல்லில் ‘*ட்*’ என்ற எழுத்தில்லாமல் ‘*சடம்*’ என இருப்பது விடுபடு பிழையாகும்.
2. செருகல் பிழை (Insertion Error): இப்பிழையில் ஒரு எழுத்து கூடுதலாக இருக்கும். எடுத்துக்காட்டாக, *சட்டடம்* என்ற பிழையான சொல்லில் ‘*ட*’ என்ற எழுத்துக் கூடுதலாக இருப்பது செருகல் பிழையாகும்.
3. பதிலீட்டுப் பிழை (Substitution Error): இப்பிழையில் ஒரு எழுத்திற்குப் பதிலாக வேறொரு எழுத்தை மாறி அமைத்து விடும். எடுத்துக்காட்டாக, *சட்டம்* என்ற சொல்லில் முதல் எழுத்து ‘*ச*’ தவறாக அடிக்கப்படும்போது அத்தவறான எழுத்து ‘ச’ ஆகவோ ‘க’ ஆகவோ ‘ப’ ஆகவோ வரலாம். இம்மாதிரியான பிழைகளைப் பதிலீட்டுப் பிழைகள் எனலாம்.
4. இடமாற்றுப் பிழை (Transposition Error): இப்பிழையில் அடுத்தடுத்த இரண்டு எழுத்துக்கள் இடம் மாற்றி இருக்கும். எடுத்துக்காட்டாக, *சட்டம்* என்ற சொல்லில் பிழையாக ‘*ட*’ என்ற எழுத்தும் ‘*ட்*’ என்ற எழுத்தும் இடம் மாறி *சடட்ம்* என்று மாற்றி அடிக்கப்படுவது இடமாற்றுப் பிழையாகும்.

2.1.3 மொழியில் இல்லாத சொற்களைக் கண்டறிந்து மாற்றுதலுக்கான செயல் திட்டம்

ஒரு மொழியில் உள்ள சொல்லினைக் கண்டறிந்து சரிசெய்வதற்கு அகராதி பயன்படுத்தப்படுகிறது. கொடுக்கப்பட்ட சொல் அகராதியில் இல்லை எனில் அச்சொல்லை மொழியில் இல்லாத சொல் எனக் கருதலாம். இதற்குத் தகுந்த கணிப்பொறிக்கான அகராதி சிறியதாக இருத்தல் நலம். மிகப் பெரிய அகராதியாக இருந்தால் அதிகமாகப் புழக்கத்தில் உள்ள எழுத்துப்பிழைச் சொற்கள் கூட அரிதாகப் பயன்படுத்தும் சரியான சொல்லாக இருக்கும்.

2.1.4 பிழை நீக்கும் வழிமுறைகள்

உரைச் செய்திகளில் உள்ள எழுத்துப் பிழைகளைக் கண்டறிந்து மாற்றுவதற்கு இருவழி முறைகள் பயன்படுத்தபடுகின்றன:

1. பெய்சியன் விதிகள் (Bayesian Rules)
2. ஓசை ஊடக மாதிரி (Noisy Channel Model)

பெய்சியன் விதிமுறைகளைப் பின்பற்றும் பொழுது எழுத்துப் பிழைகளைக் கண்டறிவதற்கான வாய்ப்புகளை ஆராய்வதற்குக் கணித முறையில் அமைந்த நிகழ்தகவுக் கோட்பாடு (Probability Theory) பயன்படுத்தப்படுகிறது. பெறும்பாலும் இம்முறைகள் தட்டச்சு செய்த உரைகளுக்குப் பயன்படுத்தப்படுகின்றன. இம்முறையை ஒளிவழி எழுத்துணரி (OCR) முறையிலும் பேச்சுப் புரிந்துகொள்ளல் ஒழுங்குமுறைகளிலும் (Voice Recognition System) உள்ளீடு செய்யப்படும் உரைகளுக்கும் பயன்படுத்தலாம்.

இரண்டாவது ஓசை ஊடக மாதிரி (Noisy Channel Model) முறையில் உள்ளீடு செய்யப்படுகின்ற உரைகள் எவ்வாறு மாற்றமடைகின்றன, மாற்றமடைவதற்கான காரணங்கள் யாவை போன்றவைகள் ஆராயப்பட்டு ஒரு மாதிரி வடிவம் (Model) தயாரிக்கப்படுகின்றது. பின்னர் இம்மாதிரி வடிவத்திற்கான எதிர்மறைச் செயல்பாடுகள் (Reverse Process) ஒருமாதிரிப் படிவமாக உருவாக்கப்பட்டுப் பிழைகள் நீக்கப்படுகின்றன. எடுத்துக்காட்டாக, தொலைபேசியில் பேசுகின்றபொழுதும் ஒலிவாங்கி வழி பேசுகின்ற பொழுதும் ஒருவர் பேசும் குரல் அமைப்புகள் மாறுபடுகின்றன. இத்தகைய கருவிகள் மூலம் கிடைக்கப்பெறுகின்ற முடிவுகள் ஒலி அமைப்பைப் பொறுத்துள்ளது; உள்ளபடியாகவே அமையாது. இந்நிலையில் அம்மாறுபாடுகள் நிகழ்வதற்கான காரணங்களை ஆராய்ந்து அதற்கு எதிர்மறையான திட்டங்களை வடிவமைக்கின்ற பொழுது இவைகள் மூலம் நிகழும் பிழைகளை நீக்கலாம்.

2.1.4.1 பெசியன் முறை (Bayesian Medhod)

ஒரு உரையிலுள்ள எழுத்துப்பிழைகளைக் கண்டறிந்து திருத்துவதற்குப் பல செயல் திட்ட முறைகள் வடிவமைக்கப்படுகின்றன. அவற்றில் பெசியன் முறை நிகழ்தகவு அடிப்படையில் (Probability Thories) அமைந்ததாகும். இம்முறையின் அடிப்படைத் தத்துவம் பின்வரும் படத்தில் விளக்கப்பட்டுள்ளது:



உரைக்கான சொற்கள் தகுந்த உள்ளீட்டுக் கருவிகள் மூலம் உள்ளீடு செய்யப்படுகின்றன. உள்ளீட்டுக் கருவி கணிப்பொறியின் விசைப்பலகையாகவோ ஒலியுணரி ஒழுங்குமுறையின் ஒலிவாங்கியாகவோ இருக்கலாம். தட்டச்சு செய்யும்பொழுது எழுத்துக்களின் சிதைவோ ஒலியுணரி ஒழுங்குமுறையில் ஒலிச் செய்தியின் சிதைவோ ஏற்பட வாய்ப்புகள் உள்ளன. மூலக்கருவிகள் மூலம் உருவாக்கப்படும் சொல் உரிய கடத்திகள் (Channels) மூலம் கடத்தப்படுகிறது. கடத்திகளில் உள்ள குறைபாடுகள் காரணமாகவும் கடத்தும் செய்திகளில் சிதைவு ஏற்படலாம். இவ்வாறாகச் சிதைந்த சொல் மறு உருவாக்கம் (Decode) செய்யப்படுகிறது. இச்சொல் மூலக்கருவி மூலம் உருவாக்கப்பட்ட சொல்லுக்கு வாய்ப்புள்ள சொல்லாகவே அமையும். இவ்வாறு சிதைந்த சொல்லுக்கு வாய்ப்புள்ள சொற்களின் பட்டியல் உருவாக்கப்படும். இப்பட்டியலிலிருந்து சிதைந்த சொற்கூறு திருத்தம் செய்யப்பெற்றுப் பிழை நீக்கிய சரியான சொல் அடையாளம் காணப்படுகிறது.

பேசியன் முறையில் எழுத்துத் திருத்திக்கான வழிமுறைகள் இருபெரும் கட்டங்களாகச் செய்யப்படுகின்றன:

1. பிழை நீங்கிச் சரியான சொல்லுக்கு வாய்ப்புள்ள சொற்பட்டியல் தயாரித்தல்

2. இச்சொற்பட்டியலில் உள்ள ஒவ்வொரு சொல்லுக்கும் மதிப்பெண் அளித்து அதிக மதிப்பெண் பெறும் சொல்லைப் பிழைத் திருத்தம் செய்யப்பட்ட சொல்லாகக் கருத வழிவகை செய்தல்.

இம்முறையில் பிழைத்திருத்தம் என்பது சொல்லில் ஏதேனும் ஒரு எழுத்து மட்டும் பிழையாக இருக்கலாம் என்ற வாதத்தின் அடிப்படையில் செயல்முறைகள் மேற்கொள்ளப்படுகின்றன. எனவே பின்வரும் வகையிலான எழுத்துப் பிழைகள் மட்டும் இம்முறையில் எடுத்துக் கொள்ளப்படுகின்றன:

- சொல்லின் ஒரு எழுத்து விடுபட்டு இருத்தல்
- சொல்லில் ஒரு எழுத்து அதிகமாக இருத்தல்
- சொல்லில் ஒரு எழுத்திற்குப் பதிலாக மற்றொரு எழுத்து தவறாக இருத்தல்
- சொல்லில் அடுத்தடுத்த இரு எழுத்துக்கள் இடம் மாறி இருத்தல்

பேசியன் முறையில் எழுத்துத் திருத்தி உருவாக்குவதற்கு முதலில் பெருந்தரவு (Corpus) ஒன்றினைக் கணிப்பொறியில் உருவாக்கிக்கொள்ள வேண்டும். இந்தப் பெருந்தரவில் மொத்தம் ‘N’ சொற்கள் இருப்பதாகக் கொள்வோம். பிழையானச் சொல்லுக்கு வாய்ப்புள்ள சொற்பட்டியல் ஒன்றை உருவாக்க வேண்டும். இப்பட்டியலை உருவாக்கும் ஒரு சொல்லில் ஒரு எழுத்தினால் மட்டும் பிழை நிகழ்வதாக எடுத்துக்கொள்ளப்படுகிறது. எனவே சொல்லில் (அ) ஒரு எழுத்து விடுபட்டு இருத்தல், (ஆ) ஒரு எழுத்து கூடுதலாக இருத்தல், (இ) ஒரு எழுத்திற்குப் பதிலாக மற்றொரு எழுத்து இருத்தல், (ஈ) அடுத்தடுத்த இரு எழுத்துக்கள் தவறாக இடம் மாறியிருத்தல் ஆகிய நான்கு மட்டும் கருத்தில் கொள்ளப்படுகின்றன. மேற்கண்ட முறைகளின்படி பிழையான சொல்லிற்கு வாய்ப்புள்ள சொற்பட்டியலில் ‘V’ சொற்கள் இருப்பதாகக் கொள்வோம். இப்பட்டியலில் ‘c’ என்ற சொல்லாக இருப்பதற்கான முன்நிகழ்வு எண்ணிக்கை பின்வரும் சூத்திரத்தின் மூலம் கணக்கிடப்படுகிறது:

$$P(c) = \frac{C(c)+0.5}{N+0.5V}$$

(V = வாய்ப்புள்ள சொல், C = விரிதரவு ‘Corpus’)

மேற்கண்ட சூத்திரம் ஒரு சொல் மட்டும் பெருந்தரவில் இடம்பெற்றுள்ள முன்நிகழ்வு (prior) மதிப்பைத் தருகிறது. இரண்டாவது கட்டநிலையில் தவறான சொல்லுக்கு வாய்ப்புள்ள சொற்களுக்கு மதிப்புகள் கண்டறியப்படுகின்றன. இதற்கு மாதிரி உரைகள் ஒரு தட்டச்சர் அல்லது பல தட்டச்சர்களைக் கொண்டு உள்ளீடு செய்யப்படுகின்றன. கெர்னிகன் மற்றும் பிறர் (Kernighan et al) ஒவ்வொரு வகைத் தனிநிலைப் பிழைக்கும் நான்கு குழப்ப அணிகளைப் (confusion matrices) பயன்படுத்தினர்.

1. Del (x, y) என்பது xy என்ற இரு எழுத்துகளுக்குப் பதிலாக x மட்டும் எத்தனை முறை தவாற அச்சடிக்கப்பட்டுள்ளது என்பதைக் குறிக்கிறது.

2. Ins (x, y) என்பது x என்ற எழுத்துக்களுக்குப் பதிலாக xy என்று தட்டச்சு செய்துள்ள எண்ணிக்கையைக் குறிக்கிறது.

3. Sub (x, y) என்பது x என்ற எழுத்திற்குப் பதிலாக எத்தனை முறை y என்ற எழுத்து தட்டச்சு செய்யப்பட்டுள்ளது என்பதைக் குறிக்கிறது.

4. Trans (x, y) என்பது xy என்ற எழுத்துக்களுக்குப் பதிலாக yx என இடமாறி அச்சடிக்கப்பட்டுள்ளது என்பதைக் குறிக்கிறது.

தவறான சொல்லுக்கு வாய்ப்புள்ள சொற்பட்டியலில் இருந்து ஏதாவது ஒரு சொல் ‘c’ என்பதை எடுத்துக்கொள்வோம். ‘c_p’ என்பது ‘c’ என்ற சொல்லில் P- இடத்தில் உள்ள எழுத்தைக் குறிப்பதாகக் கொள்வோம். இந்நிலையி ‘c’ என்ற சொல் வாய்ப்பான சொல்லாக இருப்பதற்கான மதிப்பு பின்வருமாறு கணக்கிடப்படுகிறது:

1. ஒரு எழுத்து விடப்பட்டிருக்கும் நிலையில்

$$P(t/c) = \frac{del[c_p -1,c_p]}{count\ [c_{p-1}c_p]}$$

2. ஒரு எழுத்து கூடுதலாக இருக்கும் நிலையில்

$$P(t/c) = \frac{ins\ [c_{p-1},t_p]}{count\ [c_{p-1}]}$$

3. ஒரு எழுத்திற்குப் பதிலாக மற்றொரு எழுத்து இருக்கும் நிலையில்

$$P(t/c) = \frac{trans\ [t_p,c_p]}{count\ [c_p]}$$

4. இரு எழுத்துக்கள் இடம் மாறி இருக்கும் நிலையில்

$$P(t/c) = \frac{trans[c_p,c_{p+1}]}{count\ [c_p,c_{p+1}]}$$

தவறான சொல் வாய்ப்புள்ள சொற்பட்டியலில் உள்ள ‘c’ என்ற சொல்லாக இருப்பதற்கான மதிப்பு பின்வருமாறு கணக்கிடப்படுகிறது:

$$W= P\ (t/c)*P(c)$$

இந்த சூத்திரத்தில் P(c) என்பது c என்ற சொல் பெருந்தரவில் உள்ள முன் நிகழ்வு மதிப்பைத் தருகிறது. P(t/c) என்பது தவறாகத் தட்டச்சு செய்யப்பட்ட சொல் ‘c’ என்ற சொல்லாக இருப்பதற்கு எவ்வளவு வாய்ப்புகள் உள்ளன என்ற மதிப்பைத் தருகின்றது. இவ்விரு எண்களின் பெருக்கல் தொகை, அதாவது P (t/c)*P(c) என்பது இச்சொல்லிற்கான மதிப்பெண்ணைத் தருகிறது.

மேற்கண்டவாறு தவறான சொல்லுக்கு வாய்ப்புள்ள சொற்பட்டியலில் உள்ள ஒவ்வொரு சொல்லிற்கும் மேற்கண்ட மதிப்பினைக் கண்டறிய வேண்டும். எந்தச் சொல் அதிக மதிப்பெண் பெற்றிருக்கிறதோ அந்தச் சொல் மிகச் சரியான சொல் என கருதப்படுகிறது. அதிக மதிப்பெண் பெற்றிருக்கும் சொல் பின்வரும் சூத்திரத்தின் மூலம் கூறப்பட்டுள்ளது.

$$W = \underset{c \in V}{\operatorname{argmax}} \frac{P(t/c)P(c)}{P(t)}$$

மேற்கண்டவாறு அதிக மதிப்பெண் பெற்ற சொல் இடம் பெற்ற வாக்கியத்தின் பொருளுடன் ஒப்பிடுப் பார்க்கப்படும். அது வார்த்தையின் பொருளுக்கேற்ற சொல் இல்லையெனில் அடுத்த மதிப்பெண் பெற்ற சொல் பொருத்திப்பார்க்கப்படும். இவ்வாறாகத் தவறாகத் தட்டச்சு செய்யப்பட்டுள்ள சொற்களைச் சரியான சொற்களாக மாற்றச் செயல் திட்டங்கள் உள்ளன.

இவ்வாறாக உரைகள் எழுதும்போது பிழைகளைக் கண்டறிந்து திருத்துவதற்குக் கணிப்பொறியில் பல செயல்திட்டங்கள் உருவாக்கப்படுகின்றன. அதில் மிகவும் பிரபலமானது மேற்கண்ட பேசியன் முறையாகும். இதுபோல் குறைந்த திருத்த தூரம் (minimum edit distance) என்ற மற்றொரு முறையும் பயன்படுகிறது.

2.1.4.2. குறைந்த திருத்து தூரம்

முந்தைய பகுதியில் பேசியன் வழிமுறைவரைவு எவ்வாறு திருத்தங்களைச் செய்ய முடிந்தது என்று பார்த்தோம். இருப்பினும் கெர்னிகன் மற்றும் பிறர் (Kernighan et al 1990) ஒரு சொல் ஒரு எழுத்துப்பிழை மட்டும் கொண்டிருக்கும் என்ற எளிய ஊகத்தைச் சார்ந்திருந்தனர். நாம் பன்முகப் பிழைகளைக் கையாளுவதற்குக் கூடுதல் ஆற்றலுள்ள வழிமுறை வரைவை விருப்பலாம். அம்மாதிரியான வழிமுறைவரைவைக் கோர்வை தூரத்தின் சிக்கலுக்குரிய பொதுவான தீர்வாகக் கருதலாம். ‘கோர்வை தூரம்’ என்பது எந்த அளவுக்கு இரண்டு கோர்வைகள் ஒன்றுபோல் இருக்கின்றன என்பதன் ஒரு அணியாகும். பேசியன் முறையைப் பிழைதிருத்தச் சிக்கலுக்கு அம்மாதிரியான வழிமுறைவரைவைப் பயன்படுத்தும் வழியாகப் பார்க்க இயலும்; நாம் தரப்பட்ட பிழைக்கு மிக உயர்ந்த நிகழ்தகவு கொண்டது என்ற அர்த்தத்தில் தவறுக்கு மிக அண்மையில் உள்ள சொல்லைத் தேர்வுச் சொல்லாக எடுக்கின்றோம்.

கோர்வை தூரத்தைக் கண்டுபிடிப்பதற்கு வழிமுறைவரைவுகளின் (algorithms)பிரபலமான வகுப்புகளில் ஒன்று வாக்னர் மற்றும் ஃபிஷர் (Wagner and Fischer 1974) என்பவர்களால் குறைந்த திருத்து தூரம் (minimum edit distance) என்று பெயரிடப்பட்ட (ஆனால் தனித்தனியாகப் பலரால் கண்டுபிடிக்கப்பட) ஒரு பதிப்பாகும். இரு கோர்வைகளுக்கு இடையே உள்ள குறைந்த திருத்து தூரம் என்பது ஒரு கோர்வையை மற்றொரு கோர்வையாக மாற்றத் தேவைப்படும் திருத்து நடவடிக்கைகளின்/செயன்மைகளின் (செருகல், நீக்குதல், பதிலீடுசெய்தல்) குறைந்த எண்ணிக்கையாகும். எடுத்துக்காட்டாக intension மற்றும் extension என்ற சொற்களுக்கு இடையே உள்ள இடைவெளி ஐந்து செயன்மைகளாகும்/நடவைக்கைகளாகும்; இதை மூன்று வழிகளில் உருப்படுத்தம் செய்யவியலும் ; கீழே காட்டியுள்ளது போன்ற சுவடு/தடம் (trace), வரிசையாக்கம் (alignment) அல்லது செயன்மைப் பட்டியல் (operation list) என்ற மூன்று வழிகள் உள்ளன.

சுவடு

```
 I n t e n   t   i   o   n
  / / / /    |   |   |   |
e x e c u    t   i   o   n
```

வரிசையாக்கம்

I n t e n ε t i o n

ε e x e c u t i o n

நடவடிக்கைப் பட்டியல்

I n t e n t i o n

I-ஐ நீக்கவும்

n t e n t i o n

n-ஐ e-ஆல் பதிலீடு செய்யவும்.

e t e n t i o n

t-ஐ x-ஆல் பதிலீடு செய்யவும்.

e x e n t i o n

u-ஐ செருகவும்

e x e n u t i o n

n-ஐ c-ஆல் பதிலீடு செய்யவும்

e x e c u t i o n

நாம் ஒவ்வொரு நடவடிக்கைக்கும் ஒரு குறிப்பிட்ட மதிப்பையோ (cost) சுமையையோ (weight) ஒதுக்கலாம். இரு வரிசைகளுக்கு இடையிலான லெவென்ஷ்டைன் தூரம் (Levenshtein distance) தான் மிக எளிய அளவிடும் காரணி; இதில் ஒவ்வொரு நடவடிக்கைக்கும் விலை 1 (Levenshtein 1966) ஆகும். இதன்படி intension என்பதற்கும் execution என்பதற்கும் இடையில் உள்ள தூரம் 5 ஆகும். லெவென்ஷ்டைன் தன்னுடைய அளவுமுறைக்கு மற்றொரு மாற்று பதிப்பை முன்மொழிகின்றார்; இதில் ஒவ்வொரு செருக்கலுக்கும் அல்லது நீக்கலுக்கும் மதிப்பு ஒன்று ஆகும்; பதிலீடு அனுமதிக்கப்படமாட்டாது (பதிலீட்டை ஒரு செருகலாலும் ஒரு நீக்கலாலும் உருப்படுத்தம் செய்ய இயலும் ஆகையால் ஒவ்வொரு பதிலீட்டுக்கும் மதிப்பு 2 தரப்படுகின்றது). இந்தப் பதிப்பின் பயன்பாட்டின் படி intension என்பதற்கும் execution என்பதற்கும் இடையிலுள்ள தூரம் 8 ஆகும். நாம் கூடுதல் கலவைத்தன்மையான செயல்பாடுகளால் சுமையேற்றலாம்; எடுத்துக்காட்டாக ஒவ்வொரு நடவடிக்கைக்கும் முன்னர் கூறிய குழப்ப அணிகளைப் பயன்படுத்தி ஒரு நிகழ்தகவை தருவது. இந்நேர்வில் இரண்டு கோர்வைகளுக்கு இடையிலுள்ள குறைந்த திருத்து தூரத்தைப் பற்றி பேசுவதற்குப் பதிலாக நாம் ஒரு கோர்வைக்கும் மற்றொரு கோர்வைக்கும் அதிக நிகழ்தகவு வரிசையாக்கம் பற்றி பேசுவோம். இதை நாம் செய்தால் ஒவ்வொரு மாற்றத்தின் நிகழ்தகவுகளைப் பெருக்கும் பெரிதாக்கப்பட்ட குறைந்த திருத்து தூர வழிமுறைவரைவை ஒரு தேர்வுப் பிழை தரப்படுகையில் பல் தவறு தட்டச்சின் பேசியன் சாத்தியத்தியத்தை மதிப்பிட பயன்படுத்த இயலும்.

ஆற்றலுள்ள நிரலாக்கத்தால் (dynamic programming) குறைந்த திருத்து தூரம் கணக்கிடப்படுகின்றது. ஆற்றலுள்ள நிரலாக்கம் வழிமுறைகளின் (algorithms) ஒரு வகுப்பின் பெயராகும்; இது முதலில் பெல்மான் (Bellman 1957) என்பவரால் அறிமுகப்படுத்தப்பட்டது; இது துணைச் சிக்கல்களின் திர்வுகளை ஒன்றிணைத்து சிக்கல்களைத் தீர்க்க அட்டவணை-இயக்க முறையைப் (table-driven method)

பயன்படுத்துகின்றது. இந்த வழிமுறைகளின் வகுப்பு பேச்சு மற்றும் மொழி ஆய்வுகளில் கூடுதல் பொதுவாகப் பயன்படுத்தப்படுகின்ற வழிமுறைகளை உட்படுத்தும்; எழுத்துப் பிழை திருத்ததிற்கான குறைந்த திருத்து தூர நிரல், பேச்சு புரிந்துகொள்வதிலும் இயந்திர மொழிபெயர்ப்பிலும் பயன்படுத்தப்படும் வெட்டர்பி வழிமுறை (Veterbi algorithm) மற்றும் பகுத்தாயப் (parsing) பயன்பத்தப்படும் CYK மற்றும் Early வழியமைப்பு வரைவு என்பன இவையாகும்.

2.2 இலக்கணப்பிழை திருத்தம்

இலக்கணத் திருத்தி ஒரு வாக்கியத்தின் தொடரியல் செம்மையை நிர்ணயிக்கும் அல்லது தீர்மானிக்கு இலக்கணத் திருத்தம் சொல் ஆய்விகளிலும் (word processor) தொகுப்பிகளிலும் (compilers) பெரும்பாலும் பயன்படுத்தப்படுகின்றது. தொகுப்பி போன்ற பயன்பாடுகளில் இலக்கணத் திருத்தம் எளிமையானதாகும். எனென்றால் சொற்றொகை அவற்றில் எல்லைக்குப்பட்டதாகும்; ஆனால் இயற்கை மொழிகளில் இது சவாலுக்குரியதாகும்; ஏனென்றால் அதன் சொற்றொகை எல்லைக்கு உட்பட்டதல்ல.

மொழிகளில் இலக்கணத் திருத்தத்திற்கு மூன்று நெறிமுறைகள் பயன்படுத்தப்படுகின்றன; அவையாவன:

(1) தொடரியல் அடிப்படையிலான இலக்கணத்திருத்தம் (syntax-based checking)
(2) புள்ளியியல் அடிப்படையிலான இலக்கணத்திருத்தம் (Statistics-based checking)
(3) விதி அடிப்படையிலான இலக்கணத்திருத்தம் (Rule-based checking)

2.2.1 தொடரியல் அடிப்படையிலான திருத்தம் (Syntax based checking)

தொடரியல் அடிப்படையிலான இலக்கணத் திருத்தத்தில் ஒவ்வொரு வாக்கியமும் அவற்றின் இலக்கணச் செம்மைக்காக முழுமையாகப் பகுத்தாயப்படும் (parsed). தொடரியல் பகுத்தாய்தல் பிழையானால் பனுவல்/உரை தவறானதாகக் கருத்தப்படும். ஜென்சன் (Jensan 1993) என்பவரால் உருவாக்கப்பட்ட அணுகுமுறையயில் ஒரு உரை முழுவதும் பகுத்தாயப்படும். அதாவது வாக்கியங்கள் ஆயப்பட்டு ஒவ்வொரு

வாக்கியத்திற்கும் ஒரு கிளை அமைப்பு தரப்படும். பகுத்தாய்தல் வெற்றிகரமாக நடைபெறாவிட்டால் உரையானது தவறானதாகக் கருதப்படும்.

தொடரியியல் அடிப்படையிலான அணுகுமுறையின் சிறப்பு என்னவென்றால் இலக்கணம் முழுமையாக இருந்தால் இலக்கணத் திருத்தலும் முழுமையாக அமையும். அதாவது இலக்கணத் திருத்தி தவறு எவ்வளவு தெளிவற்று இருந்தாலும் எந்தத் தவறான வாக்கியத்தையும் கண்டுபிடிக்கும். துரதிஷ்ட வசமாக இலக்கணத் திருத்தி தவறான வாக்கியங்களை மட்டும் தெரிந்துகொள்ளும்; அது பயன்படுத்துவோருக்குச் சரியாக என்ன சிக்கல் என்பதைக் கூற இயலாது. இதற்காகக் கூடுதல் விதிகள் தேவைப்படும்; அதுவும் தவறான வாக்கியங்களைப் பகுத்துக்குறிக்கும். ஒரு வாக்கியத்தை அம்மாதிரியான கூடுதல் விதியால் மட்டும் பகுத்துக்குறிக்க இயலுமென்றால் அவ்வாக்கியம் தவறானதாகும். இந்த நுட்பம் கட்டுப்பாடு நெகிழ்வு (constraint relaxation) எனப்படும்.

எனினும், தொடரியல் சார்ந்த அணுகுமுறையில் ஒரு முக்கிய பிரச்சனை உள்ளது: அதற்கு ஒருவர் சரிபார்க்க விரும்பும் அனைத்து வகையான பனுவல்களையும் உள்ளடக்கிய ஒரு முழுமையான இலக்கணம் தேவை. பல இலக்கண கோட்பாடுகள் இருந்தாலும் இன்னும் வலுவான பரந்த உள்ளடக்கப்பரப்புப் பாகுப்பாய்வி (robust broad-coverage parser) பலரறிய எதுவும் கிடைக்கவில்லை. மேலும் பாகுப்பாய்விகள் இயற்கையான பொருண்மை மயக்கத்தால் பாதிக்கப்படுகின்றன; இதன் காரணமாகச் சரியன வாக்கியங்களுக்கும் வழக்கமாக ஒன்றுக்கும் மேற்பட்ட விளைவுகள் திரும்பக்கிடைக்கின்றன.

2.2.2. **புள்ளியியல் அடிப்படையிலான திருத்தம்** (Statistical based checking)

புள்ளியியல் அடிப்படையிலான இலக்கணத்திருத்தத்தில் சொல்வகைப்பாடு அடையாள (POS tag) வரிசைகள் ஒரு அடையாளப்படுத்தப்பட்ட தரவுத்தொகுதியிலிருந்து உருவாக்கப்படும்; இவ்வரிசைகளின் நிகழ்வெண் அதாவது நிகழ்தகவு/சாத்தியக்கூறு (probability) குறித்துக்கொள்ளப்படும். சொல்வகை அடையாளப்படுத்தப்பட்ட உரை/பனுவல் சில நிகழ்வெண்ணைக் காட்டிலும் குறைவான நிகழ்வெண்ணில் சொல்வகைப்பாடு வரிசைகளைக் கொண்டிருந்தால் அப்பனுவல் பிழையானதாகக் கருத்தப்படும். புள்ளியியல் அடிப்படை அணுகுமுறை முக்கியமாக

அடையாளப்படுத்தப்பட்ட பயிற்சி தரவுத்தொகுதியிலிருந்து (tagged training corpus) விதிகளைக் கற்றுக்கொள்ளும். விதி அடிப்படை அணுகுமுறையில் விதிகள் கையால் எழுத்தப்படவேண்டும் என்பது தவிர அணுகுமுறை புள்ளியியல் அணுகுமுறையுடன் ஒற்றுமை உடையதாகும்.

இருப்பினும் ஆங்கிலத்திற்கு மிகவும் பயன்படுத்தப்படுகின்ற இலக்கணத்திருத்திகளில் ஒன்றான மைகிரோசாஃப்ட் ஆபீஸ் சூட் (Microsoft Office Suit) இலக்கணத் திருத்தியும் எதிர்கருத்திற்கு (controversy) அப்பாற்பட்டதல்ல. தற்போது இலக்கணத் திருத்தி உருவாக்கம் மிகவும் எளிமையான செயல் அல்ல என்பதை நிரூபிக்கின்றது.

இது அட்வல் (Attwell1987) என்பவரால் உருவாக்கப்பட்ட அணுகுமுறையில் சொல்வகைப்பாட்டுக்கு அடையாளப்படுத்தப்பட்ட தரவுத் தொகுதி சொல்வகைப்பாடு அடையாளப்படுத்தப்பட்ட கோர்வைகளின் பட்டியலை உருவாக்கப் பயன்படுத்தப்படும். சில கோர்வைகள் பொதுவானதாகும்.

எ.கா.

அடைகொளி அடை + பெயரடை+ பெயர் = The old man.

பிற எல்லாவற்றிலும் நேர்வதில்லை.

எ.கா.

அடைகொளி அடை + பெயரடை

தரவுத் தொகுதியில் அடிக்கடி நேரும் கோர்வைகளைப் பிற உரைகளில் சரியானதாகக் கருத இயலும்; அரிதாக நேரும் கோர்வைகள் தவறானதாகக் கருதப்படும்.

புள்ளியியல் அணுகுமுறையில் என்கிராம் பகுப்பாய்வை (n-gram analysis) பயன்படுத்தி ஒரு வாக்கியத்தின் சாத்தியக்கூறை அளக்கவியலும். எடுத்துக்காட்டாக, He is playing என்ற வாக்கியத்தின் பைகிராம் சத்தியக்கூறு பின்வருமாறு அமையும்.

P(“He is playing.”) = P(He|<start>)*P(is|he)*P(playing|is)*P(.|playing)

இம்மூன்று சொற்களில் ஏதேனும் ஒன்று பயிற்சித் தரவுத்தொகுதியில் இல்லாவிடில், வாக்கியத்தின் சாத்தியக்கூறு, பெருக்குத்தொகையின் காரணமாகப் பூஜியமாக மாறும்.

நாம் இப்புள்ளியியல் நெறிமுறையில் சொற்களைக் கருத்திற்கொண்டால் மொழியின் எல்லாச் சொற்களையும் கொண்டுள்ள பெரிய தரவுத்தொகுதி நமக்குத் தேவைப்படும்.

இச்சிக்கலைத் தீர்ப்பதற்காக நாம் தனிச்சொற்களுக்குப் பதிலாகச் சொல்வகைப்பாட்டு அடையாளங்களைப் (POS tags) பயன்படுத்த இயலும். வேறுபாடு என்னவென்றால், முன்னர் நாம் எச்சொற்கள் ஒரு சொல்லுக்குப்பின் வரும் கூடுதல் சாத்தியக்கூறைக் கொண்டிருக்கின்றன என்பதைப் பரிசோதித்தோம்; தற்போது எச்சொல்வகை அடையாளங்கள் ஒரு சொல்வகைப்பாட்டு அடையாளத்திற்குப்பின் வரும் கூடுதல் சாத்தியக்கூறைக் கொண்டிருக்கின்றன என்று பரிசோதிப்போம். நாம் சொல்வகைப்பாடு அடையாளங்களைப் பயன்படுத்தினால், சொற்கள் மாறிகளாகும் (variables).

“He is playing.” என்ற முந்தைய எடுத்துக்காட்டு வாக்கியத்தை எடுத்துக்கொள்ளவும். சொல்வகைப்பாடு அடையாளப்படுத்தலுக்குப் பிறகு அவ்வாக்கியம் He/pps is/bez playing/vbg./.” என மாறும். இப்போது வாக்கியத்தின் சாத்தியக்கூறை கணக்கிட நாம் சொல்வகைப்பாட்டு வரிசையைப் பயன்படுத்த இயலும்.

P(pps bez vbg.) = P(pps|<start>)*P(bez|pps)*P(vbg|bez)*P(.|vbg)

நாம் கூறுகிற இலக்கண திருத்தி பின்வருமாறு செயலாற்றுகின்றது.

1. வாக்கியத்தின் ஒவ்வொரு சொல்லுக்கும் ஒரு அடையாளம் தரவும்.
2. அடையாள வரிசையின் (tag sequence) சாத்தியக்கூறை தீர்மானிக்க என்கிராம் பகுப்பாய்வைப் பயன்படுத்தவும் (இதில் என்=3 அதாவது டிரை கிராம்)
3. சாத்தியக்கூறு சில நிகழ்வெண்ணிற்கு மேல் இருந்தால் வாக்கியம் இலக்கண அடிப்படையில் சரியானது என்று கருதப்படும். நமது மாதிரியில் சாத்தியக்கூறு பூஜியத்திற்கு அதிகமாக இருந்தால் வாக்கியம் சரியானது என்று கருதப்படும். இரண்டோ அதற்கு மேற்பட்ட அடுத்துவருகிற அடையாளங்கள் பொருந்த இயலவில்லை என்றால் (அதாவது இயைபின்றி இருந்தால்) வரிசையின் சாத்தியக்கூறு பூஜியமாகும்.

முதலில் நமக்கு ஒரு சொல்வகை அடையாளப்படுத்தி (POS tagger) தேவை. அது தானியக்கமாகச் சொற்களை அடையாளப்படுத்தும்; அல்லது நாம் வாக்கியத்தின் சொற்களைக் கையால் அடையாளப்படுத்தவேண்டும். பின்னர் அடையாள வரிசையின் சாத்தியக்கூறைத் தீர்மானிக்க (இரண்டு முந்தைய அடையாளங்களைப் பார்க்கும்) டிரை கிராம் மாதிரியைப் பயன்படுத்தவும்; இறுதியாக அவ்வடையாள வரிசையின் சாத்தியக்கூறின் அடிப்படையில் இலக்கணச் செம்மையின் தீர்மானத்தைச் செய்யவும்/எடுக்கவும். எடுத்துக்காட்டாக பிரவுண் தரவுத்தொகுதி (Brown Corpus) மற்றும் பிரில்லின் சொல்வகைப்பாடு அடையாளப்படுத்தியை (Brill's tagger), “He saw the book on the table” என்ற வாக்கியத்தின் கணக்கீடுகள் பின்வருவனவாகும்:

He/pps saw/vbd the/at book/nn on/in the/at table/nn ./.

P(pps|None None) = 0.0635486169593

P(vbd|None pps) = 0.213047910296

P(at|pps vbd) = 0.166456494352

P(nn|vbd at) = 0.48308668061

P(in|at nn) = 0.362738953306

P(at|nn in) = 0.350967938144

P(nn| in at = 0.44004695623

P(.|at nn) = 0.0847696646819

அடையாள வரிசையின் சாத்தியக்கூறு = 5.16478478489e-09

இலக்கணத்திருத்தியின் முடிவு = இவ்வாக்கியம் சாத்திய அடிபடையில் சரியானதாகும்.

நாம் உடன்பாட்டு இயைபு இல்லாத “He have the book I want” என்ற வாக்கியத்தை ட்ரைகிராம் அடிப்படையில் ஆய்ந்தால் இலக்கணத் திருத்தி பின்வரும் கணக்கீடுகளைத் தரும்.

He/pps have/hv the/at book/nn I/ppss want/vb./.

P(pps|none none) = 0.0635486169593

P(hv|none pps) = 0.0

P(at|pps hv) = 0.0

P(nn|hv at) = 0.0491712707182

P(ppss|at nn) = 0.00493575681605

P(vb|nn ppss) = 0.293785310734

P(.|ppss vb) = 0.0361445783133

அடையாள வரிசையின் சாத்தியக்கூறு =0.0

இலக்கணத்திருத்தியின் முடிவு = தவறானது அல்லது கண்டுபிடிப்பது சாத்தியமில்லை

புள்ளியியல் சார்ந்த பாகுபடுத்திகள் அவற்றின் முடிவுகளை விளக்குவது கடினம் என்ற இடரைக் கொண்டிருக்கின்றது: ஒரு தவறான எச்சரிக்கைப் பிழை இருந்தால் பயன்பாட்டாளர் குறிப்பிட்ட பிழைச் செய்தி இல்லாத காரணத்தால் ஏன் அவரது உள்ளீடு தவறாகக் கருதப்படுகிறது என்று ஆச்சரியப்படுவர். உருவாக்குபவர் கூட ஒழுங்குமுறையின் தீர்ப்பைப் புரிந்துகொள்ள ஒழுங்குமுறைக்குப் பயிற்சி அளிக்கப்பட்ட தரவுத்தொகுதியை அணுக வேண்டிவரும். மற்றொரு சிக்கல் யாராவது ஒருவர் பொதுவல்லாத மற்றும் தவறான கட்டுகளைப் பொதுவல்லாத ஆனால் சரியான கட்டுகளிலிருந்து பிரிக்கும் தொடக்க நிலையை நிறுவவேண்டும். இந்தச் செயல் 0-வுக்கும் 100-க்கும் இடையில் ஏதாவது மதிப்பை நிறுவும் பயன்பாட்டாளருக்குத் தரப்படவேண்டும். தொடக்க நிலையின் கருத்து (நடை மற்றும் கட்டப்பட்ட இறுதி நேர்வுகள் தவிர) வாக்கியங்கள் வழக்கமாகச் சரி அல்லது தவறு என்ற புலனுணர்வுடன் உண்மையில் இணங்கவேண்டும் என்பதில்லை.

2.2.3. **விதி அடிப்படையிலான திருத்தம் (Rule based checking)**

இது டேனியன் ராபர் மற்றும் பிறரால் (Denial Naber et al) பயன்படுத்தப்பட்ட அணுகுமுறையில் குறைந்தது சொல்வகைப்பாட்டுக்காக அடையாளப் படுத்தப்பட்ட

உரை ஒரு குழும விதிகளாகப் பொருத்தம் பார்க்கப்படும். இந்த அணுகுமுறை புள்ளியியல் அடிப்படையிலான அணுகுமுறைக்கு ஒற்றுமையுள்ளது; ஆனால் உரைகள் மனிதனால் உருவாக்கப்பட்டதாகும்.

பிற அணுகுமுறைகளில் கூறப்பட்ட சிக்கல்கள் காரணமாகக் கண்டிப்பான விதி அடிப்படையிலான ஒழுங்குமுறை இங்கு உருவாக்கப்படும். தொடரியல் அடிப்படையிலான பிழை திருத்திகளைப் போல அல்லாமல் விதி அடிப்படையிலான பிழை திருத்தி ஒரு போதும் முழுமையாக இருப்பதில்லை; அதாவது அதில் கண்டுபிடிக்க இயலாத பிழைகள் இருந்துகொண்டே இருக்கும். இருப்பினும் அதற்கு பல அனுகூலங்கள் உண்டு:

- ஒரு வாக்கியம் சரிபார்க்க முழுமையாக இருக்கத் தேவையில்லை; மாறாக மென்பொருள் பனுவலை அது தட்டச்சு செய்யப்படும் போதே சரிபார்த்து உடனடியான பதில் கருத்தைத் தரவியலும்.
- ஒவ்வொரு விதியும் வெளிப்பாட்டு விளக்கம் கொண்டிருப்பதால் அதற்கு உருவங்கொடுப்பது எளிது மற்றும் தனித்தனியாக முடுக்கவோ அணைக்கவோ இயலும்.
- அது பயனுள்ள கருத்துரைகளுடன் இலக்கண விதிகளை விளக்கியே விளக்கமான பிழைச் செய்திகளைத் தர இயலும்.
- விதி ஒழுங்கமைப்பு குறைந்தது பல எளிய ஆனால் பொதுவான பிழை நேர்வுகளை எளிதில் புரிந்துகொள்ள இயலுமாகையால் அதை அதன் பயன்பாட்டாளர்களுக்கு எளிதாக நீட்சி செய்யவியலும்
- அதைப் படிப்படியாக ஒரு விதியில் தொடங்கி பின்னர் விதிக்கு விதி நீட்சி செய்து உருவாக்கவியலும்.

2.3 நடை மற்றும் இலக்கணத்திருத்தி

நடை மற்றும் இலக்கண திருத்தி (style and grammar checker) ஒரு பனுவலை/உரையை ஏற்கும் பின்னர் சாத்தியமான தவறுகளைத் திருப்பி அனுப்பும். தவறுகளை/பிழைகளைக் கண்டுபிடிப்பதற்காக பனுவலின் ஒவ்வொரு சொல்லுக்கும்

அதன் சொல்வகைப்பாடு மதிப்பிடப்படும்/தரப்படும்; ஒவ்வொரு வாக்கியமும் தொடர்கூறுகளாக (chunkers) பிரிக்கப்படும். எடுத்துகாட்டாகப் பெயர்த்தொடர்கள் பின்னர் பனுவல் பிழை திருத்தியின் முன்வரையறை செய்யப்பட்ட பிழை விதிகளுடன் ஒப்பிட்டுப்பார்க்கப்படும். ஒரு விதி பொருந்தினால் பொருந்தும் இடத்தில் உரையில் பிழை இருப்பதாகக் ஊகிக்கப்படும்/கருதப்படும். விதிகள் பிழைகளைச் சொற்கள், சொல்வகைப்பாட்டு அடையாளங்கள், தொடர்கூறுகள் இவற்றின் அமைப்பொழுங்கை (pattern) விளக்கும். ஒவ்வொரு விதியும் பிழைக்கு விளக்கத்தை உட்படுத்தும்; அது பயன்படுத்துபவர்க்குக் காட்டப்படும்.

அமைப்பொழுங்கின் விதி அடிப்படையிலான அணுகுமுறை பயன்படுத்துபவர்கள் தங்களது விதிகளை உருவாக்கிக்கொள்ள இயலச்செய்கின்ற அளவுக்கு எளிமையானதாகும்; பல எடுத்துக்காட்டான பிழைகளைக் கண்டுகொள்வதற்குப் போதுமான அளவு திறன் வாய்ந்ததாகும். விதிகள் எளிய XML வடிவமைப்பில் வெளிப்படுத்தப்படுகின்றது; இது பிழைகளை விளக்குவதுடன் பயன்படக்கூடிய பிழைச் செய்திகளையும் எடுத்துக்காட்டு வாக்கியங்களையும் கொண்டிருக்கின்றது.

2.3.1. கோட்பாட்டுப்பின்னணி

எழுத்துப்பிழை திருத்தம் பயனுள்ளதாகக் கருதப்பட்ட அதே காரணத்திற்காக நடை மற்றும் இலக்கணப்பிழை திருத்தமும் பயனுள்ளதாகும். இது மக்கள் பிழையில்லாமல் ஆவணங்களை எழுதுவதற்குப் பயன்படுகின்றது. நடை மற்றும் இலக்கணப்பிழை திருத்தி பயனுள்ளதாக அமையச் சில தேவைகளைப் பூர்த்திசெய்யவேண்டும்.

- அது விரைவாக இருக்கவேண்டும்; அதாவது இடையீடுப் பயன்பாட்டிற்குப் போதுமான அளவு விரைவாக இருக்கவேண்டும்.
- அது இருக்கின்ற சொல்லாய்வியுடன் நன்றாக இணைக்கப் பட்டிருக்கவேண்டும்.
- சரியான வாக்கியங்களைப் பெரும்பான்மையும் தவறாகக் காட்டக்கூடாது.
- சொந்தத் தேவைகளுக்காக அதை மாற்றிக்கொள்வது சாத்தியப்படவேண்டும்.
- இறுதியாக அது முழுமையாக இருக்கவேண்டும்; அதாவது அது பனுவலில் உள்ள பெரும்பான்மையான பிழைகளைக் கண்டுபிடிக்கவேணடும்.

எழுத்துப்பனுவலில் தோன்றும் பிழைகளைப் பல்வேறு பட்ட வழிகளில் வகைப்படுத்த இயலும்.

எழுத்துப் பிழைகள்: இது எழுத்துப்பிழை திருத்தியால் கண்டுபிடிக்க இயலும் பிழையாகும் என வரையறை விளக்கம் செய்யப்படுகின்றது. இலக்கணப் பிழைதிருத்திகள் அறியப்பட்ட சொற்களின் பெரிய பட்டியலுடன் சொற்களை ஒப்பிட்டுப்பார்க்கும். சொல் அந்தப் பட்டியலில் இல்லாவிட்டால் அது பிழையானது என்று கருதப்படும். ஒற்றுமையுள்ள (சரியான) சொற்கள் பதில் சொற்களாகத் தரப்படும். எடுத்துக்காட்டாக *Gemran என்பது German எனத் திருத்தப்படும்.

இலக்கணப் பிழைகள்: ஒரு மொழியின் இலக்கண விதிகளுடன் ஒத்துப்போகாத ஒரு வாக்கியத்தை விளைவிக்கும் பிழை இலக்கணப் பிழையாகும். எழுத்துப்பிழை திருத்தம் போல அல்லாமல் இலக்கணப்பிழை திருத்தம் சூழல் ததகவலைப் பயன்படுத்த வேண்டியது அவசியமாகும்.

*Harry Potter bigger then than Titanic?

இப்பிழை தட்டச்சால் ஏற்பட்டதா, எழுதுபவரின் மனதில் then என்பதற்கும் than என்பதற்கும் இடையில் ஏற்பட்ட தவறான புரிந்துகொள்கையால் ஏற்பட்டதா என்பதைத் தீர்மானிக்க இயலாது. then என்பதைப் பயன்படுத்துவது தவறு என்று தெளிவாகத் தெரிவதால் இது இலக்கணப் பிழையாகும்.

இலக்கணப் பிழைகளை அமைப்புசார் பிழை மற்றும் அமைப்புசாரா பிழை என்று பகுக்கலாம். அமைப்பு சார்ந்த பிழைகளைச் சொற்களைச் செருகியோ நீக்கியோ ஒன்றோ இரண்டோ சொற்களை நகர்த்தியோ மட்டுமே திருத்த இயலும். அமைப்புசாராத பிழைகளை இருக்கிற சொல்லை வேற்று சொல்லால் இடம்பெயர்த்துத் திருத்தம் செய்ய இயலும்.

நடைப்பிழைகள்: இயல்பற்ற/அசாதாராண சொற்களும் சிக்கலான வாக்கிய அமைப்புகளும் உரையை/பனுவலைப் புரிந்துகொள்வதைக் கடினமாக்குகின்றது; இது யாராலும் விரும்பப்படுவதில்லை. இந்நேர்வுகள் நடைப் பிழைகள் என்று கருதப்படும்.

இலக்கணப்பிழைகளைப் போல் அல்லாமல் நடைப் பிழைகள் அதிகமாகச் சூழலைச் சார்ந்தும் பனுவல்/உரை வகையைச் சார்ந்தும் அமையும். எடுத்துக்காட்டாக, சொந்த நண்பர்களுக்கிடையிலான மின்னஞ்சல் கருதுப்பரிமாற்றம் மொழியின் படைப்பாக்கப் பயன்பாட்டை அனுமதிக்கின்றது; அனால் தொழில் நுட்ப ஆணவப்படுத்தல் பொருண்மை மயக்கத்தால் பாதிக்கலாகாது. அமைப்பு உறுதிப்பாடு (configurability) இலக்கணப் பிழை திருத்தங்களை விட நடை பிழை திருத்தத்திற்குக் கூடுதல் முக்கியத்துவம் வாய்ந்ததாகும்.

எடுத்துக்காட்டாக, பினவரும் வாக்கியத்தை எடுத்துக்கொள்ளவும்.

> But it [a human reason] quickly discover that, in this way, its labours must remain ever incomplete, because new questions never cease to present themselves, and thus it finds itself compelled to have recourse to principles which transcend the region of experiences, while they are regarded by common sense without distrust.

இவ்வாக்கியம் உண்மையான காரணத்தால் காந்தின் விமர்சனத் திறனாய்விலிருந்து (Kant's Critique) விளைந்ததாகும். இது 48 சொல் நீளமுள்ளதாகும்; பலரும் இது புரிந்துகொள்வதற்குக் கடினமானது என்று ஒத்துக்கொள்வர். இதற்குக் காரணம் அதன் நீளம், (transcend போன்ற) கடினமான சொற்றொகை, இரட்டை எதிர்மறையின் (double negation) பயன்பாடு. இக்காலகட்டத்தில் புரிந்துகொள்வதற்கு எளிமையான ஆணவத்தின் தேவையுள்ள சூழலில், இவ்வாக்கியத்தில் நடைச் சிக்கல் உள்ளதாகக் கருதலாம்.

பொருண்மைப் பிழைகள்: நடைப் பிழையோ இலக்கணப்பிழையோ எழுத்துப் பிழையோ இல்லாத தவறான தகவல் உள்ள ஒரு வாக்கியத்தில் பொருண்மைப் பிழை உள்ளதாகக் கருதலாம். பொருண்மைப் பிழைகளைக் கண்டுபிடிக்க மிகப் பரந்த உலகறிவு (world knowledge) தேவையாதலால், இப்பிழைகளை வழக்கமாகக் கண்டுபிடிக்க இயலாது. எடுத்துக்காட்டாக, "MySQL is a great editor for programming!" என்ற வாக்கியம் உண்மையோ பொய்யோ அல்ல; MySQL என்பது ஒரு எடிட்டர் அல்ல ஒரு டேட்டா பேஸ் ஆகும் என்பதால் இது பொருளைத் தரவில்லை. பரந்த உலகறிவு இல்லாமல் இதைத்

தெரிந்துகொள்ள இயலாது. உலகறிவு சூழலின் ஒரு வடிவாகும்; ஆனால் இன்றைய காலகட்டத்தில் இது மென்பொருளின் புரிந்துகொள்ளலுக்கு அப்பாற்பட்டதாகும்.

பிழைகளுக்கும் (errors) தவறுகளுக்கும் (mistakes) வேறுபாடு பாராட்டவேண்டியதில்லை. தவறாகப் கருதப்படுபவைகளும் மோசமாக எழுதப்படுபவைகளும் பிழைகளாகும். இலக்கணம் (அல்லது தொடரியல்) சரியான வாக்கியங்கள் எவ்வாறு இருக்கவேண்டும் என்பதை விளக்கும் விதிகளின் ஒழுங்கமைப்பாகும் (system of rules). எவ்வாறோ விதிகள் மக்கள் மனதில் இருக்கின்றது; இதன் காரணமாகப் பெரும்பாலான மக்கள் ஒரு வாக்கியம் தவறானதா இல்லையா என்பதை எளிதில் தீர்மானிக்க இயலும். உள்ளுணர்வு / மன உணர்வு (intuition) அடிப்படையில் சரியானது/தவறானது என்ற தீர்மானத்தைக் கடினமாக்கும் எல்லையை ஒட்டிவரும் நேர்வுகளைக் காண்பது சாத்தியமாகும். பிழையான செய்திக்குப் பதிலாக விளக்க எச்சரிக்கைச் செய்தியைக் (descriptive warning message) காட்டி ஒரு மென்பொருள் இந்நேர்வுகளைக் கையாளும். இலக்கண விதி சில சரியான வாக்கியங்களுக்கும் பொருந்துவது தெரிந்தால், எச்சரிக்கைச் செய்திகளும் நடைப் பிழைகளுக்கும் இலக்கணப் பிழைகளுக்கும் பயனுள்ளவையாக அமையும்.

2.3.2 சொல்வகை அடையாளப்படுத்தல் (Parts-of-speech tagging)

சொல்வகை அடையாளப்படுத்தல் (parts-of-speech tagging) ஒவ்வொரு சொல்லுக்கும் அதன் சொல்வகையை அடையாளப் படுத்துவதாகும். எந்தச் சொல்வகை அடையாளங்கள் உள்ளனன என்று அறுதியிட்டு விளக்கப்படவில்லை; அனால் மிகப் பொதுவானவைகள் பெயர், வினை, வரையறை அடை (அடைகொளிஅடை), பெயரடை, வினையடை என்பனவாகும். பெயர்களை ஒருமை, பன்மை என்றும் வினைகளை இறந்தகாலம், நிகழ்காலம், எதிர்காலம் என்றும் இது போன்று மேலும் பகுக்கலாம்.

அதிகமாகச் சொல்வகை அடையாளங்கள் இருந்தால் அது கூடுதல் சிக்கலுக்கு வழிவகுக்கும்; ஏனென்றால் பல சொற்களுக்கு சூழல் அடிப்படையில் வேறுபட்ட சொல்வகைபாடுகள் இருக்கக்கூடும். ஆங்கிலத்தில் பல சொற்களைப் பெயராகவும்

வினையாகவும் பயன்படுத்த இயலும்; எடுத்துக்காட்டாக house என்பது 'கட்டிடம்' என்றும் 'பாதுகாப்பு அளி' என்றும் பொருள்படும். அனைத்துச் சொற்களிலும் 11.5% தான் சொல்வகைப்பாட்டைப் பொறுத்தவரையில் மயக்கம் உடையன; இவை அடிக்கடி நேரும் சொற்கள் ஆகையால் ஒரு பனுவலில் 40% சொற்கள் மயக்கம் உடையன. சொல்வகைப்பாடு அடையாளப்படுத்துதல் ஒரு மாதிரி சொல்மயக்க நீக்கச் சிக்கலாகும்.

சூழலைக் கருத்தில் கொள்ளாமல் தரப்பட்ட ஒரு சொல்லுக்கு அடிக்கடி நேரும் சொல்வகைப்பாட்டு அடையாளத்தைத் தேர்ந்தெடுத்து ஒருவர் 91% சொற்களுக்கு அவற்றின் சரியான அடையாளத்தைத் தரவியலும். பெரிய தரவுத் தொகுதியின் புள்ளியியல் ஆய்வின் அடிப்படையிலான பெரும்பான்மையான அடையாளப்படுத்திகள் 95-97% துல்லியத்தைக் காட்டுகின்றன (Brill, 1992). கட்டுப்பாட்டு இலக்கணம் (constraint grammar) 99% துல்லியத்திற்குக் கூடுதல் உரிமைகோருகின்றன (Karlsson, 1995). இவ்விழுக்காடுகளைப் புரிந்துகொள்ளுவதில் நாம் விளிப்புணர்வுடன் இருக்கவேண்டும். சில அடையாளப்படுத்திகள் சில மயக்கம் தரும் நேர்வுகளை விட்டுவிட்டு ஒரு தரப்பட்ட சூழலில் ஒரு சொல்லுக்கு ஒன்றிற்கும் மேற்பட்ட சொல்வகைப்பாட்டு அடையாளத்தை திருப்பி அனுப்பும். திருப்பி அனுப்பப்பட்ட சொல்வகை அடையாளங்களில் குறைந்தது ஒன்று சரியானது என்ற சாத்தியங்கள் ஒரு சொல்வகை அடையாளம் திருப்பி அனுப்பப்படுவதை விடக் கூடுதலாகும். வேறுவிதமாகக் கூறினால் ஒருவர் மீட்டழைப்புக்கும் (recall) துல்லியத்திற்கும் (precision) இடையில் வேறுபாடு கருதவேண்டும்.

2.3.3 தொடர் பகுத்தாய்தல் (Phrase chunking)

தொடர் பகுத்தாய்தல் சொல்வகைப்பாடு அடையாளப்படுத்தலுக்கும் ஒரு முழு இலக்கணப் பகுப்பாய்வுவுக்கும் இடையே அமைந்துள்ளது: சொல்வகைப்படு அடையாளப்படுத்தல் சொல்நிலையில் மட்டுமே செயல்பாடு செய்யும் போது மற்றும் இலக்கணப் பகுப்பாய்வு (பகுத்தாய்தல்) வாக்கியத்திற்கு கிளையமைப்பு கட்டமைக்கும் போது, தொடர் பகுத்தாய்தல் ஒரு வாக்கியத்திலுள்ள சொல் வரிசைகளுக்கு ஒரு அடையாளத்தை ஒதுக்குகிறது. பெயர்த்தொடர் (பெதொ) மற்றும் வினைத்தொடர் (விதொ) என்பன எடுத்துகாட்டான பகுப்புகளாகும். பெயர்த்தொடர் வரையறை

அடைகள், பெயரடைகள், பெயர்கள், மாற்றுப்பெயர்கள் ஆகியனவற்றைக் கொண்டிருக்கும். வினைத்தொடர் ஒரு தனிநிலை வினை அல்லது துணை வினை மற்றும் வினையெச்சம் ஆகியவற்றைக் கொண்டிருக்க இயலும். எடுத்துக்காட்டாக the dog, the big dog, the big brown dog என்பன பெயர்த்தொடர்களுக்கு எடுத்துக்காட்டுகளாகும். பெயர்த்தொடர்களின் பட்டியல் எல்லையற்று நீளுமாகையால் கோட்பாடு அடிப்படையில் பெயர்த்தொடர் எல்லையற்று நீளும். இருப்பினும் இங்கு பெயர்த்தொடர் என்று அழைக்கப்படுவது ஒரு எடுத்துக்காட்டு தான்; சொல்வகை அடையாளப்படுத்தல் போல ஒவ்வொன்றும் அதனுடைய பகுப்புப் பெயர்களையும் அவற்றின் பொருண்மைகளையும் உருவாக்க இயலும். ஒரு பகுப்பாய்வியால் கண்டுகொள்ளப்படும் பகுப்புகள் பெயர்த்தொடரையும் வினைத்தொடரையும் மட்டும் வைத்துக்கொண்டு முழுப் பனுவலையும் உட்படுத்தத் தேவையில்லை; இது சாத்தியமும் இல்லை.

சொல்வகைப்பாடு அடையாளப்படுத்தல் சொல்லில் செயல்படுவது போன்று தொடர் பகுத்தாய்தல் சொல்வகைப்பாடு அடையாளப்படுத்தப்பட்ட பனுவலில் செயல்படும்: எந்தச் சொல்வகைப்பாடு அடையாள வரிசை எந்தத் தொடர் பகுப்பாய்வுகளைக் கட்டமைக்கும் என்ற கையால் எழுதிய விதிகள் உள்ளன அல்லது நிகழ்வு சாத்தியத் தொடர்க்கூறு பகுப்பாய்வி சொல்வகைப்பாடு அடையாளப்படுத்தப்பட்ட மற்றும் தொடர்கூறு பகுத்தாயப்பட்ட பனுவல்களில் பயிற்சி அளிக்கப்படும். தொடர்கூறு பகுத்தாய்தல் ஒரு சொல்வகைப்பாடு அடையாளப்படுத்தப்பட பனுவலை வேண்டுவதால், அதன் துல்லியம் பயன்படுத்தப்பட்ட சொல்வகை அடையாளப்படுத்தியை விடச் சிறப்பாக இராது. தொடர்கூறு பகுத்தாய்தலுக்குப் பட்டியலிடப்பட்டுள்ள சிறப்பான தொடர்கூறு பகுப்பாய்வி ஒரு துல்லியத்தை அடைகின்றது மற்றும் 94% திருப்பி அழைத்தலை அடைகின்றது; இது ஒரு சராசரி சொல்வகை அடையாளப்படுத்தி எய்துவதைவிடக் குறைவாகும்.

2.3.4 இலக்கணப் பிழைகள்

எளிய மொழியான ஆங்கிலத்திற்குக் கூட இலக்கண விதிகள் விரிவானவை. இங்கு விளக்கப்பட்டுள்ள இலக்கண விதிகள் இவ்விதிகளை மீறும் தரவுத்தொகுதியிலிருந்து எடுக்கப்பட்ட வாக்கியங்கள் அடிப்படையிலானது.

2.3.4.1 எழுவாய்-வினை உடன்பாடு

ஆங்கிலத்தில் எழுவாயும் வினையும் எண் மற்றும் இடம் அடிப்படையில் உடன்பாடு காட்டும். எடுத்துகாட்டாக *They is my favourite Candian authors என்ற வாக்கியத்தில் எழுவாயும் வினையும் எண்ணில் வேறுபடுகின்றன (they = பலர்பால், is = ஒருமை. * He am running for president, என்பதில் எழுவாயும் வினையும் இடம் அடிப்படையில் உடன்படவில்லை (he = படர்க்கை இடம், am = to be என்பதன் தன்மையிடம்)

இது எளிய நேர்வாகும். சொல்வகைப்பாட்டு அடையாளங்களைக் கொண்ட டோக்கன்களின் வரிசைகளாக பனுவலை எடுத்துக்கொள்ளும் விதி அடிப்படையிலான திருத்தியின் பார்வையை எடுத்துக்கொண்டால் பல சிறப்பான நேர்வுகள் உள்ளன:

1. எழுவாயும் வினையும் பிரிக்கப்பட்டுள்ளன, அதாவது வினை எழுவாய்க்கு நேரடியாகப் பின்னர் வருவதில்லை: *The characters in Shakespear's Twelth Night lives in a world that has been turned upside-down.
2. எழுவாய் ஒரு கூட்டுநிலை எழுவாயாக இருக்கவியலும்: *Chrisie and Prin is characters form Laurence's The diviners.
3. நூலின் தலைப்புகள் எல்லாம் ஒருமை: *Salman Rushdie's Midnight's children are my favourite novel.

2.3.4.2 குறிப்பில்லா அடை மற்றும் தொடர்ந்துவரும் சொல்

குறிப்பில்லா அடையைத் தொடர்ந்துவரும் சொல்லின் உச்சரிப்பு உயிர் ஒலிப்பு முதலில் வரும் படி இருந்தால் a அடைக்குப் பதிலாக an என்ற அடை பயன்படுத்தப்படவேண்டும்.

எ.கா.

A test, a car, a long talk

An idea, an uninteresting speech, an earth quake

2.3.4.3 தொடுவினா

பேச்சு மொழியில் தொடுவினா (tag question) ஒரு கூற்றுவாக்கியத்திற்கு உறுதியைப் பெறுவதற்குப் பயன்படுத்தப்படும். இது ஆங்கிலத்தில் துணைவினையின் எதிர்மறை வடிவையும் வாக்கியத்தின் இறுதியுடன் வாக்கியத்தின் எழுவாயைச் சேர்ப்பதாலும் கட்டமைக்கப்படுகின்றது. எடுத்துக்காட்டாக, It's warm today என்பது it is warm today, isn't it? என வரும். தொடுவினாக்கள் மின் அஞ்சலில் பயன்படுத்தப்படுகின்றது. ஜெர்மானிய தாய்மொழி பேசுபவர்களுக்கு ஆங்கித்திலுள்ள தொடுவினா அமைப்பைப் புரிந்து கொள்வது கடினம். எனவே அவர்கள் ஆங்கிலத் தொடுவினா வடிவை ஜெர்மானிய மொழிக் கட்டமைப்பின் படி தவறாகப் பயன்படுத்துவர்.

2.3.4.4 பிற பிழைகள்

பெரும்பாலான பிற பிழைகள் தொழிநுட்ப அடிப்படையில் இலக்கணப் பிழைகள்; அவைகள் தட்டச்சால் உருவாக்கப்படுகின்றது. பெரும்பாலும் பிழைகள் அது பனுவலைத் திருத்தும்போது சில சொற்களை விட்டுவிடுவதால் ஏற்படுகின்றது.

எ.கா.

*Someone suggested said that it worked for him after he updated to Kernel 2.4.20.

இவ்வாக்கியத்திற்கு உரியவர் said என்பதை suggested என்பதால் இடம்பெயர்க்க விரும்புகிறார்; ஆனால் said என்பதை நீக்குவதற்கு மறந்துவிடுகிறார்.

பெரும்பாலும் ஒன்றுபோல் இருக்கின்ற சொற்கள் கலக்கப்படுகின்றன; மற்றும் எழுத்தாளர் தட்டச்சர் பிழை செய்தாரா அல்லது அவருக்கு வேறுபாடு தெரியவில்லையா என்று கூற இயலாது.

2.3.4.5 வாக்கிய எல்லை கண்டுபிடிப்பு

இலக்கணத்தன்மை வாக்கியங்களைக் குறிப்பிடும். பின்வரும் இரண்டு வாக்கியங்களின் இலக்கணத்தில் பிழை உள்ளது; Peter என்ற பெயருக்கும் she என்ற பதிலீட்டுப்பெயருக்கும் உடன்பாடு இல்லாததால் அவைகளுக்கிடையில் இலக்கணப் பிழை உள்ளது.

Peter leaves his house. She feels good.

இம்மாதிரியான பிழைகளைக் கையாண்டு திருத்துவது கடினம். இது பொருண்மை நிலையில் வரும் பிழையாகும். மனிதர்களுக்கு இருக்கும் இம்மாதிரியான உலக அறிவைக் கணிப்பொறிக்குப் பகிர்ந்து தருவது கடினம்.

சுருக்கச் சொற்கள், எண்கள், அயற்கூற்று என்பன கணிப்பொறிகளுக்குக் கடினமானது. வாக்கர் மற்றும் பிறர் (Walker et al, 2001) வாக்கிய எல்லை கண்டுபிடிப்புக்கு மூன்று அணுகுமுறைகளை மதிப்பீடு செய்கின்றார். அவர்களின் கவனக்குவிப்பு தங்கள் ஒழுங்கமைப்பில் வாக்கிய எல்லைக் கண்டுபிடிப்பு தேவைப்படுகின்ற தானியங்கு மொழிபெயர்ப்பு ஒழுங்குமுறைகளாகும். இம்மூன்று அணுகுமுறைகள் பின்வருவனவாகும்:

- சொற்களையும் நிறுத்தற்குறிகளையும் தவிர உயர்நிலை விளக்கங்களைப் பயன்படுத்தாமல் மொழிபெயர்ப்பு ஒழுங்குமுறையில் நேரடியாக நடைமுறைப் படுத்தல்: இந்த ஒழுங்குமுறை சுருக்க அகராதியைப் (abrivation lexicon) பயன்படுத்துகின்றது. இதை நடைமுறைப்படுத்துவது சீரான வெளிப்பாடுகளால் (regular expressions) ஊக்கப்படுத்தப்பட்டுள்ளது; ஆனால் எல்லாம் தரப்பட்டுள்ள ஒரு வழியமைப்பு மொழியில் நேரடியாக நடைமுறைப்படுத்தப்பட்டுள்ளது.
- சீரான வெளிப்பாடுகளாக வாக்கியங்களின் விதி அடிப்படையிலான உருப்படுத்தம். சீரான வெளிப்பாடுகள் கூற்றுவழியில் தேவையான அறிவை குறியாக்கம் செய்ய அனுமதிக்கின்றன. இது வெளிப்படையாகக் கூறப்படாவிட்டாலும் இந்நெறி ஒரு சுருக்கச்சொல் அகராதியையும் பயன்படுத்துவதாகத் தெரிகின்றது.
- அடையாளப்படுத்தப்பட்ட தரவுத்தொகுதியில் பயிற்சியளிக்கப்பட்ட இயந்திரம் கற்கும் வழிமுறை வரைவின் பயன்பாடு. இவ்வழிமுறை வரைவு தலைப்பெழுத்தாக்கம் மற்றும் சுருக்கச்சொல் அகரதி போன்ற சாத்தியமான வாக்கிய எல்லையின் பல பண்புக்கூறுகளை அளவிடுகின்றது.

வாக்கர் மற்றும் பிறரின் மதிப்பீடு இயந்திரம் கற்றல் வழிமுறைவரைவு சுமார் 98% துல்லியத்தையும் திரும்பப் பெறுதலையம் தருகின்றது என்று காட்டுகின்றது.

எனவே மிக நல்ல எல்லை கண்டுபிடித்தலுடன் கூடிய ஒரு நடை மற்றும் இலக்கண பிழைதிருத்தி வாக்கிய எல்லைகளில் கிட்டத்தட்ட 2% பிழை விகிதத்துடன் செயல்படும். இது தவறாகப் தொடர்கூறு பகுக்கப்பட்ட வாக்கியங்களின் மேல் செயல் புரியும்படி எதிர்பார்க்கப்படுகின்ற பகுத்துக்குறிப்பானின் மோசமான சிக்கலாகும்; ஏனென்றால் தவறாகச் சேர்க்கப்பட்ட வாக்கிய எல்லையும் தவறப்பட்ட வாக்கிய எல்லையும் பகுத்துக்குறிக்க இயலாத உள்ளீட்டிற்குக் கொண்டு செல்லும். விதி அடிப்படையிலான ஒழுங்குமுறைக்கு இது குறைந்த சிக்கல் உள்ளது: வாக்கிய எல்லைகளை வெளிப்படையாகக் குறிப்பிடும் விதிகளே பாதிக்கப்படும். இவ்விதிகள் வாக்கிய எல்லைகள் தவறாகச் சேர்க்கப்படும் போது தூண்டப்படும் மற்றும் வாக்கியக் கண்டுபிடிப்பால் வாக்கிய எல்லை தவறப்படும் போது அவை தூண்டப்படாது இருக்கும்.

2.3.5 கட்டுப்படுத்தப்பட்ட மொழித் திருத்தம்

கட்டுப்படுத்தப்பட்ட மொழி என்பது மொழியை எளிமையானதாகவும் புரிவதற்கு எளிதாகவும் ஆக்குவதை நோக்கமாகக் கொண்ட விதிகளினால் கட்டுப்படுத்தப்பட்ட இயற்கை மொழியாகும். மற்றொரு வழியில் கூறினால் கட்டுப்படுத்தப்பட்ட மொழி இயற்கைமொழியின் ஒரு துணைநிலை மொழியாகும். கட்டுப்படுத்தப்பட்ட மொழியை உருவாக்குவதன் முக்கியமான நிலைப்பாடு எல்லா மொழியியல் நிலைகளிலும் மயக்கத்தை விலக்குவதாகும். மயக்கமில்லாத மொழிகளுக்கு பொதுவான இயற்கைமொழிகளோடு ஒப்பிடுகையில் இரண்டு முக்கியமான அனுகூலங்கள் உள்ளன.

- குறிப்பாக இயற்கைமொழியின் இயல்பான பேசுபவர்களாக இல்லதிருக்கையில், அது புரிவதற்கு எளிதாகும். இயல்பான மொழி பேசுபவர்களுக்கும் தவறாகப் புரிதல் மிகவும் குறைக்கப்படும். இது ஆவணங்களைப் படிப்பதை எளிதாக்குவதுடன் சில வகை ஆவணங்களில் மிக

முக்கியமானதாகும்; எடுத்துக்காட்டாக வானவூர்தி இயந்திரங்களைப் பாதுகாப்பது பற்றி விளக்கும் ஆவணங்கள்.

- இது கணிப்பொறியால் பகுத்துக்குறிப்பதற்கு எளிதாகும்; எனவே தானியக்கமாக மொழிபெயர்ப்பதற்கு எளிதாகும். இது மொழிபெயர்ப்புச் செயன்முறையில் மிகச் சிக்கனமானதாகும்; இது மனித முயற்சியால் செய்யப்படும் போது அதிக செலவீனமுள்ளதாகும்.

கட்டுப்படுத்தப்பட்ட மொழியின் எடுத்துக்காட்டான பயன்பாடுகள் பின்வருமாறு அமையும்:

- **சொல் கட்டுப்பாடுகள்**: எடுத்துக்காட்டாக try என்ற சொல்லை வினையாக மட்டும் பயன்படுத்தவும்; பெயராகப் பயன்படுத்தவேண்டாம். இது பொருண்மையியல் கட்டுப்பாட்டை குறிப்பாக உணர்த்தும்; ஏனென்றால் அது தொடரியல் கட்டுப்பாடாக வெளிப்படுத்தப்பட்டால் அது try என்ற பெயரை நேரிடும் போது ஒரு பிழைச் செய்தியை தூண்டும் ஒரு விதியுடன் காணப்படும்.

- **இலக்கணக் கட்டுப்பாடுகள்**: எடுத்துக்காட்டாக வினையை ஏவல் வடிவில் எழுதவும்.

  Remove oil and grease with a degreasing agent.

  Oil and grease are to be removed with a degreasing agent.

- **பொருண்மை கட்டுப்பாடுகள்**: எடுத்துக்காட்டாக noise என்ற சொல்லை " unwanted sound" என்ற பொருளில் மட்டும் பயன்படுத்தவும்.

- **நடை கட்டுப்பாடுகள்**: ஒரு விதி வாக்கியங்களை 20 சொற்களுக்குள் வைத்துக்கொள்ளவும் ஒரு வாக்கியத்திற்கு ஒரு தலைப்பு எனக் கொள்ளவும் கோரலாம்.

2.3.6 நடை திருத்தம்

இயற்கைமொழிகளின் இலக்கணம் சரியான வாக்கியங்களுக்கும் தவறான வாக்கியங்களுக்கும் இடையிலான வேறுபாட்டைத் தெளிவாகத் தருகின்றது. இது மொழி நடைக்குப் பொருந்தாது. ஒவ்வொ எழுத்தாளரும் எந்த நடை எந்த இடத்தில் விருப்பத்தேர்வு செய்யப்படவேண்டும் என்று தமது சொந்தத் தீர்மானங்களை எடுக்கின்றனர். பலரால் ஒரு ஆவணம் இணைந்து உருவாக்கப்படும் போது இது இணக்கம்/இயைபு இல்லாத நடைக்குக் கொண்டுசெல்லும். இம்மாதிரியான சூழல்களில் நடை திருத்தம் மிகப் பயனுள்ளதாகும். நடையைப் பற்றிய கட்டுப்பாடுகளும் விதிகளும் விரிந்த எல்லையிற் படும் விசயங்களை உட்படுத்தும்.

- **சொற்களின் விருப்பத்தேர்வு**: எல்லோரும் புரிந்துகொள்வதற்குக் கடினமாகையால் வேற்றுமொழியின் சொற்கள் நிராகரிக்கப்படலாம். நடை திருத்தி அச்சொல்லை அதே பொருள்கொண்ட இயல் மொழிச் சொல்லால் இடம்பெயர்க்கும் படி பரிந்துரைக்கலாம்.

- **எளிமை**: புரிந்துகொள்ளும் எளிய வாக்கிய அமைப்புகளுக்குக் கொண்டு செல்லும் படி வாக்கியங்களின் நீளம் குறைக்கப்படலாம்.

- **நிறுத்தற்குறிகள்**: நிறுத்தற்குறிகளுக்குப் பின்னர் இடம் விடப்படலாம்; ஆனால் நிறுத்தற்குறிகளுக்கு முன்னர் இடம் விடப்படுவதில்லை.

- நாட்களும் நேரங்களும் எண்களும் ஒரு குறிப்பிட்ட நடை வடிவிலேயே எழுத்தப்படவேண்டும். எடுத்துக்காட்டாக 5/7/1999 என்று எழுதுவது மயக்கம் தரும்; மாறாக அது 1999-07-05 என எழுதப்படவேண்டும்.

- Don’t போன்ற சுருக்கவடிவங்கள் நிராகரிகரிக்கப்பட்டு do not என்பதால் இடம்பெயர்க்கப்படலாம்.

- ஒரு குறிப்பிட்ட பெயரை அடிக்கடி பயன்படுத்தும் பனுவல் வழக்கமற்றதாகத் தோன்றலாம். பதிலாக பதிலீடு பெயர்கள் அல்லது ஒருபொருள் பன்மொழிகள் அந்தப் பெயரை குறிப்பிட பயன்படுத்தப்படவேண்டும். தொழில்நுட்ப

ஆணவங்களில் தெளிவுக்கும் மயக்கமின்மைகளுக்கும் முக்கியத்துவம் தரப்படவேண்டும்.

2.4. முடிவுரை

எழுத்து மற்றும் இலக்கணத் திருத்தத்திற்கான சாத்தியங்களையும் வழிமுறைகளையும் கோட்பாடுகளையும் கொள்கைகளையும் பற்றி இவ்வியலில் விளக்கப்பட்டுள்ளது. இதன் அடிப்படையில் தமிழுக்கான எழுத்து மற்றும் இலக்கணத் திருத்தி உருவாக்குவதற்கான சாத்தியங்கள் ஆயப்படும்.

இயல் 3

தமிழில் எழுத்துப்பிழைதிருத்தியும் இலக்கணப் பிழைத்திருத்தியும் உருவாக்குவதற்கான முன்னேற்பாடுகள்

3.0 முன்னுரை

தமிழ் எழுத்துப்பிழை திருத்தி உருவாக்கத்திற்கும் இலக்கணப்பிழை திருத்தி உருவாக்கத்திற்கும் வேண்டி முன்னேற்பாடுகளாகத் தமிழ் எழுத்தியியல் ஒலியனியல் அமைப்பமையும் இலக்கண அமைப்பையும் எழுத்து வருகைமுறை விதிகளையும் (similar to phonotactic rules) மற்றும் உருபன் வருகைமுறை விதிகளையும் (morphotactic rules) கணிப்பொறி அறியச்செய்தல் முக்கியமானதாகும்; மேலும் தமிழ்த் தொடரமைப்பு விதிகளும் கணிப்பொறிக்குத் தரப்படவேண்டும். இவற்றைக் கருத்திற்கொண்டு இவ்வியல் எழுத்தப்பட்டுள்ளது.

3.1 தமிழ் எழுத்தியல்

தமிழ் எழுத்தியல் தமிழ் ஒலியனியல் அடிப்படையில் அமைந்துள்ளது என்றாலும் ஒலியனியலில் கண்டுபிடிக்கப்பட்டுள்ள எல்லாத் தமிழ் ஒலியன்களுக்கும் தமிழில் எழுத்து இருகின்றது என்றோ எல்லா ஒற்றெழுத்துகளும் ஒலியகள் என்றோ கூற இயலாதது. தமிழ் எழுத்தியல் ஆங்கிலத்தைப் போன்று நெடுங்கணக்கு சார்ந்து (alphabetical) அல்ல. தமிழ் எழுத்துக்கள் எல்லாம் அடிப்படையில் அசை எழுத்துக்கள் (syllabic) அதாவது உயிரெழுத்தும் மெய்யெழுத்தும் சேர்ந்த உயிர்மெய் எழுத்துக்களாகும். பெரும்பாலான நேர்வுகளில் அசை எழுத்தை உருவாக்க ஒரு துணைநிலை குறியீடு மெய்யெழுத்துக்கு முன்னரோ பின்னரோ சேர்க்கப்படுகின்றது; சில நேர்வுகளில் மெய்யெழுத்து புதிய வடிவை அடையும் படி மாற்றப்பட்டுள்ளது. இவ்வாறு அசை எழுத்துக்கள் 12 உயிர் எழுத்துக்கள் மற்றும் 18 மெய்யெழுத்துக்கள் இவற்றின் ஒன்றிணைப்பால் ஆக்கப்பட்டுள்ளன; இதன் படி 216 உயிர்மெய் எழுத்துக்கள் உருவாக்கப்படுகின்றன. மரபிலக்கணப் படி உயிரெழுத்துக்கள் 12-உம் மெய் எழுத்துக்கள் 18-உம் உயிர் மெய்யெழுத்துக்கள் 216-உம் ஆயுத எழுத்து 1-உம் சேர்ந்தது ஆக மொத்தம் 247 எழுத்துக்கள் உள்ளன. இது தவிர 6 கிரந்த எழுத்துக்களும் பயன்படுத்தப்பட்டு வருகின்றது. தமிழில் ஆங்கிலம் போன்று பெரிய எழுத்து (capital

letter) சின்ன எழுத்து (small letter) என்ற வேறுபாடு கிடையாது. கீழே தமிழ் எழுத்துக்களின் (உயிரெழுத்துக்கள், மெய்யெழுத்துக்கள், உயிர்மெய் எழுத்துக்கள்) அட்டவணை தரப்பட்டுள்ளது.

தமிழ் உயிர் எழுத்து, மெய்யெழுத்து மற்றும் உயிர்மெய் எழுத்துக்களின் அட்டவணை

		a	aa	i	ii	u	uu	e	ee	ai	o	oo	au
		அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஒ	ஓ	ஔ
k	க்	க	கா	கி	கீ	கு	கூ	கெ	கே	கை	கொ	கோ	கௌ
c	ச்	ச	சா	சி	சீ	சு	சூ	செ	சே	சை	சொ	சோ	சௌ
T	ட்	ட	டா	டி	டீ	டு	டூ	டெ	டே	டை	டொ	டோ	டௌ
t	த்	த	தா	தி	தீ	து	தூ	தெ	தே	தை	தொ	தோ	தௌ
p	ப்	ப	பா	பி	பீ	பு	பூ	பெ	பே	பை	பொ	போ	பௌ
R	ற்	ற	றா	றி	றீ	று	றூ	றெ	றே	றை	றொ	றோ	றௌ
ng	ங்	ங	ஙா	ஙி	ஙீ	ஙு	ஙூ	ஙெ	ஙே	ஙை	ஙொ	ஙோ	ஙௌ
Nj	ஞ்	ஞ	ஞா	ஞி	ஞீ	ஞு	ஞூ	ஞெ	ஞே	ஞை	ஞொ	ஞோ	ஞௌ
N	ண்	ண	ணா	ணி	ணீ	ணு	ணூ	ணெ	ணே	ணை	ணொ	ணோ	ணௌ
nd	ந்	ந	நா	நி	நீ	நு	நூ	நெ	நே	நை	நொ	நோ	நௌ
M	ம்	ம	மா	மி	மீ	மு	மூ	மெ	மே	மை	மொ	மோ	மௌ
N	ன்	ன	னா	னி	னீ	னு	னூ	னெ	னே	னை	னொ	னோ	னௌ
Y	ய்	ய	யா	யி	யீ	யு	யூ	யெ	யே	யை	யொ	யோ	யௌ
R	ர்	ர	ரா	ரி	ரீ	ரு	ரூ	ரெ	ரே	ரை	ரொ	ரோ	ரௌ
L	ல்	ல	லா	லி	லீ	லு	லூ	லெ	லே	லை	லொ	லோ	லௌ
V	வ்	வ	வா	வி	வீ	வு	வூ	வெ	வே	வை	வொ	வோ	வௌ
zh	ழ்	ழ	ழா	ழி	ழீ	ழு	ழூ	ழெ	ழே	ழை	ழொ	ழோ	ழௌ
L	ள்	ள	ளா	ளி	ளீ	ளு	ளூ	ளெ	ளே	ளை	ளொ	ளோ	ளௌ
J	ஜ்	ஜ	ஜா	ஜி	ஜீ	ஜு	ஜூ	ஜெ	ஜே	ஜை	ஜொ	ஜோ	ஜௌ
S	ஸ்	ஸ	ஸா	ஸி	ஸீ	ஸு	ஸூ	ஸெ	ஸே	ஸை	ஸொ	ஸோ	ஸௌ
sh	ஷ்	ஷ	ஷா	ஷி	ஷீ	ஷு	ஷூ	ஷெ	ஷே	ஷை	ஷொ	ஷோ	ஷௌ
H	ஹ்	ஹ	ஹா	ஹி	ஹீ	ஹு	ஹூ	ஹெ	ஹே	ஹை	ஹொ	ஹோ	ஹௌ

3.2 தமிழ் ஒலியனியல் அமைப்பு

தமிழ் ஒலியனியல் வளைநா மெய்யொலிகள் மற்றும் பன்னிலை உச்சரிப்புகள் இவற்றால் பண்பாக்கம் செய்யப்பட்டுள்ளது. ஒலியனியல் அடிப்படையில் தமிழ் குரல் மெய் ஒலியன்களுக்கும் குரலிலா மெய் ஒலியன்களுக்கும் இடையில் வேறுபாடு பாராட்டுவதில்லை; ஒலியியல் அடிப்படையில் சொல்லில் மெய்யொலியின் இருப்பிடத்தைப் பொறுத்து குரல் பண்புக்கூறு தரப்படும். தமிழ் ஒலியனியல் சில மெய்மயக்கங்களை (consonant clusters) அனுமதிக்கின்றது, ஆனால் மெய் மயக்கங்கள் சொல்லின் தொடக்கத்தில் அனுமதிக்கப்படுவதில்லை. தமிழ் மரபிலக்கணத்தார் தமிழ் ஒலியன்களை உயிர்கள், மெய்கள் மற்றும் ஆயுதம் என்று வகைபடுத்துகின்றார்கள்.

3.2.1 உயிரொலியன்கள்

உயிரொலியன்கள் உயிரெழுத்து என்று தமிழ் மரபிலக்கணங்களால் அழைக்கப்படுகின்றன. உயிரொலின்கள் குறில், நெடில் மற்று ஈருயிர் என்று வகைப்படுத்தப்படுகின்றன. இலக்கணப் பனுவல்கள் குறில்களை நெடில்களுடன் வைத்து விளக்கினாலும் நெட்டுயிர்கள் குற்றுயிர்களைவிட இருமடங்கு நீளமானவை.

ஓருயிர்கள்

	இதழ்விரி			இதழ்குவி
	முன்		நடு	பின்
மூடிய/மேல்	ஈ	இ		உ ஊ
இடை	ஏ	எ		ஓ ஒ
திறந்த/கீழ்			ஆ அ	

ஈருயிர்கள்: ஐ, ஔ

3..2. 2. மெய்யொலியன்கள்

மெய்யொலியன்கள் மெய்யெழுத்து என்று தமிழ் மரபிலக்கணக்கணங்களால் அழைக்கப்படும். மெய்யொலிகள் ஒன்றில் ஆறுவீதம் மூன்று வகைபாடுகளாக வகைபடுத்தப்பட்டுள்ளது: வல்லினம், மெல்லினம், இடையினம். தமிழில் கட்டுப்பாட்டிற்கு உட்பட்ட மெய்மயக்கங்கள் உள்ளன. தமிழ் மெய்யொலிகளின் அட்டவணை கீழே தரப்பட்டுள்ளது.

	ஈரிதழ்	பல்லிதழ்	பல்	அண்பல்	வளைநா	அண்ணம்	பின்னண்ணம்	குரல்வளை
அடைப்பொலி	ப்		த்	ற்	ட்	ச் ஜ்	க்	
மூக்கொலி	ம்		ந்	ன்	ண்	ஞ்	ங்	ஹ்
உரசொலி				ஸ்	ஷ்			
மருங்கொலி				ல்	ள்	ழ்		
வருடொலி				ர்				
உயிர்ப்போலி		வ்				ய்		

தமிழ் மொழிக்கு உயிர்ப்பொலி இல்லை. தமிழ் எழுத்தில் குரலொலிக்கும் குரலிலா ஒலிக்கும் வேறுபாடு இல்லை என்றாலும், அவைகள் பேச்சு மொழியில் மாற்றொலிகளாக துணைநிலை வருகை தந்து வேறுபாடு காட்டும்.

தமிழ் மொழியில் 12 உயிர் ஒலிகள் உள்ள. இவற்றில் ஐந்து குறில் ஒலிகளாகவும், அவற்றுடன் தொடர்புடைய ஐந்து நெடில்களாகவும் இரண்டு கூட்டொலிகளாகவும் உள்ளன. *அ*, *இ*, *உ*, *எ*, *ஒ* ஆகிய ஐந்தும் குறில்; இவற்றின் நெடில்களாக *ஆ*, *ஈ*, *ஊ*, *ஏ*, *ஓ* ஆகிய ஐந்து உள்ளன. (*அ*+*இ* >) *ஐ*, (*அ*+*உ* >) *ஔ* ஆகிய இரண்டும் கூட்டொலிகள். *இ* மேல் இதழ் விரி முன்னுயிர் குறில் எனவும் *ஈ* மேல் இதழ் விரி முன்னுயிர் நெடில் எனவும் *எ* இடை இதழ் விரி முன்னுயிர் குறில் எனவும் *ஏ* இடை இதழ் விரி முன்னுயிர் நெடில் எனவும் *அ* கீழ் இதழ் விரி நடு உயிர் குறில் எனவும் *ஆ* கீழ் இதழ் விரி நடு உயிர் நெடில் எனவும் *உ* மேல் இதழ் குவி பின்னுயிர் குறில் எனவும் *ஊ* மேல் இதழ் குவி பின்னுயிர் நெடில் எனவும் *ஒ* இடை

இதழ் குவி பின்னுயிர் குறில் எனவும் *ஓ* இடை இதழ் குவி பின்னுயிர் நெடில் எனவும் பெயர்பெறும். 18 மெய்யொலிகளில் 6 வெடிப்பொலிகள் (Plosives) அல்லது அடைப்பொலிகள் (Stops) (*க்*, *ச்*, *ட்*, *த்*, *ப்*, ற்), 6 மூக்கொலிகள் (*ங்*, *ஞ்*, *ண்*, *ந்*, *ம்*, *ன்*), 3 மருங்கொலிகள் (*ல்*, *ள்*, *ழ்*), 1 வருடொலி (*ர்*), 2 அரை உயிர்கள் (ய்,வ்) எனப்படும். நுனி நா வளைந்து அண்ணத்தில் படும்போதும் உண்டாகும் 4 ஒலிகள் நாவளை ஒலிகள் (*ட்*, *ண்*, *ள்*, *ழ்*) (retroflexes) எனப்படும். நுனி நாக்கு அண்ணத்தினை ஒட்டி ஆடலும் உண்டு. ஆட விடும் போது பிறக்கும் ஒலி ஆடொலி (*ற்*) எனப்படும் (trill). நுனி நா முன் அண்ணத்தில் ஒரு தடவை சிறிது நேரம் தட்டி வருடுமாறு செய்யும் போது பிறக்கும் ஒலி வருடொலி (flap) (*ர்*) எனப்படும். வருடுவதற்குப் பதில் தொடர்ந்து சிறிது நேரம் இருப்பதால் உருவாவது தொடரொலி (*ழ்*) ஆகும்.

பிழை ஆய்வுக்கு பயன்படுத்தபபடும் யூனிக்கோடு எழுத்துரு அடிப்படையிலான
தமிழ் ஒலியன்களின் அல்லது எழுத்துக்களின் அட்டவணை கீழே தரப்பட்டுள்ளது.

	Bilabial	Labio-dental	Dental	Alveolar	Retroflex	Palatal	Velar
Plosive	ப் p		த் t	ற் R	ட் T	ச் c	க் k
Nasal	ம் m		w	ன் n	ண் N	ஞ் nj	ங் ng
Lateral Continuant				ல் l	ள் L	ழ் z	
Flap				ர் r			
Semivowel		வ் v				ய் y	

3.2.3 உயிர்ஒலிகளின் வருகை முறை

அ

அ முதல், இடை, கடை ஆகிய மூன்று இடங்களிலும் வரும்; மொழி முதலில் தனித்த மெய்களுக்கு முன் *அ* வரும்.

எ.கா.

அக்கரம், அங்கை, அசல், அஞ்ச, அடக்கி, அண்டம், அதன், அந்த, அப்படி, அம்பு, அயல், அரசு, அலைந்து, அவள், அழுதல், அளவு, அறம், அன்பு

அ மொழி இடையில் தனித்த மெய்களுக்குப் பின் *ங்*, *ஞ்*, ஆகிய மெய்களைத் தவிர பிற மெய்களுடன் வரும்.

எ.கா.

அகம், அசல், கூடம், தணல், வேதன், நகர், பராபரம், அகலாமல், நாயகன், பரம், பலம், பாவம், விளக்கு, இறந்தவர், வானம்

அ மொழி இடையில் எல்லா தனித்த மெய்களுக்கு முன்னரும் வரும்.

எ.கா.

அடக்கி, அடங்க, அரசு, கஞ்சா, கடல், கண்ணு, கதவு, அனந்தன், கப்பல், கம்பம், கயறு, சர்க்கரை, சலனம், சவம், பழக்கம், வள்ளல், அகற்றிவிடு, அகன்று

அ மொழி இறுதியில் எல்லாத் தனித்த மெய்களுக்குப் பின்னரும் வரும்.

எ.கா.

அடங்க, கொஞ்ச, அறிந்திட, உண்ண, எட்டாத, எழுப்ப, ஓட்டிய, உணர, பல, சிவ, வாழ, மெல்ல, வென்ற, என்ன

ஆ

ஆ முதல், இடை, கடை ஆகிய மூன்று இடங்களிலும் வரும்; ஆ மொழி முதலில் எல்லா மெய்யொலிகளுக்கு முன்னரும் வரும்.

எ.கா.

ஆகமம், ஆங்காரம், ஆசான், ஆஞ்சனேயர், ஆடி, ஆண்டி, ஆத்தாள், ஆந்தை, ஆபரணம், ஆமிசம், ஆய்ந்து, ஆரியன், ஆலம், ஆவி, ஆழ, ஆள, ஆறு, ஆனால்

ஆ மொழி இடையில் ங தவிர எல்லா மெய்யொலிகளுக்குப் பின்னரும் வரும்.

எ.கா.

காண், சாகாத, ஞானம், தேடாது, வீணாய், வீழ்ந்தார், நாகம், பாட, மாங்காய், முடியாது, ராசா, வல்லார், வாக்கு, வாழ்க்கை, ஆளாதே, மாறாது, முன்னால்

ஆ மொழி இடையில் எல்லா மெய்யொலிகளுக்கு முன்னரும் வரும்.

எ.கா.

சாகாமல், அட்டாங்கயோகம், உல்லாசம்,காஞ்சனம், காடு, காண், காத்திருந்தேன், காந்தம், சாபம், சூடாமல், தாய், அஞ்சார், காலம், காவல், தாழும், தாள், சாற்று, தீண்டினாய்

ஆ இறுதியில் *ங்*, *ஞ்*, *ந்* தவிர எல்லா மெய்யொலிகளுக்குப் பின்னரும் வரும்.

எ.கா.

ஆக, கஞ்சா, எட்டா, காணா, பிதா, அப்பா, சும்மா, மாயா, வாரா, எல்லா, துறவா, வாழா, மாறா, நடவா

இ

இ மொழி முதல், இடை, கடை ஆகிய மூன்று இடங்களிலும் வரும். *இ* மொழி முதலில் எல்லா மெய்களுக்கு முன்னரும் வரும்.

எ.கா.

இக்கணம், இங்கே, இசைந்து, இஞ்சி, இடம், இணை, இதழ், இந்த, இப்படியே, இம்மென்று, இயற்கை, இரண்டு, இலவு, இவள், இழந்து, இளப்பம், இறந்தார், இன்னம்

இ மொழி இடையில் *ங்*, *ஞ்* தவிர எல்லா மெய்களுக்குப் பின்னரும் வரும்.

எ.கா.

ஆக்கிய, எச்சில், ஓட்டினேன், கண்ணியர், கதிர், நிறம், சாப்பிட, அமிழ், கயிறு, எரித்தாள், சலியாமல், சேவித்து, தொழில், அளித்தவர், அறிந்தோம், சூனியம்

இ மொழி இடையில் எல்லா மெய்களுக்கு முன்னரும் வரும்.

எ.கா.

அம்பிகை, இங்கே, எண்திசை, வஞ்சியர், பிடித்து, பிணி, விதி, பிரிந்தவர், துதிப்பாய், பத்திமார்கள், பத்தியம், மந்திரம், மயில், வடிவு, விழி, விளக்கம், மின்னல்

இ மொழி இறுதியில் ங், ஞ், ந், ய் தவிர எல்லா மெய்களுக்குப் பின்னரும் வரும்.

எ.கா.

அடக்கி, அரிசி, இப்படி, ஆணை, முச்சந்தி, விரும்பு, உமி, வலம்புரி, காதலி, குலாவி, கோழி, தள்ளி, அறி, தனி

ஈ

ஈ மொழி முதல், கடை ஆகிய மூன்று இடங்களிலும் வரும். *ஈ* முதலில் க், *ச்*, *ட்*, *த்*, *ந்*, *ர்*, *ற்*,*ன்* என்ற மெய்களுக்கு முன்னர் வரும்.

எ.கா.

ஈக்கள், ஈசர், ஈட்டி, ஈது, ஈந்து, ஈராறு, ஈற்று, ஈன்று

ஈ மொழி இடையில் *ங்*, *ஞ்*, *ல்*, *ழ்*, *ள்*, *ற்* என்ற எல்லா மெய்கள் தவிர பிற எல்லா மெய்களுக்குப் பின்னரும் வரும்.

எ.கா.

கீழ், *சீர்*, காண்டீபன், தண்ணீர், *தீ*ங்கான, நீர்க்குமிழி, பீற்றல், மீன், கனிமொழிச்சியீர், நீரீழிவு, வீடு, பன்னீர்

ஈ மொழி இடையில் *ஞ்*, *ந்*, *ய்*, *ம்* தவிர எல்லா மெய்களுக்கு முன்னரும் வரும்.

எ.கா.

*தீ*க்கனல், *தீ*ங்கான, மாரீசன், மீட்டல், வீணாசை, *கீ*தம், *தீ*பம், தின்பீராகில், நீலமடி, சீவன், நீரீழிவு, நீள், பிற்றல், மீன்

ஈ மொழி இறுதியில் *த்*, *ந்* என்ற மெய்களுக்குப் பின்னர் வரும்.

எ.கா.

தீ, *நீ*

உ

உ மொழி முதலிலும் இடையிலும் கடையிலும் வரும். *உ* மொழி முதலில் *ஞ்* தவிர எல்லா மெய்களுக்கு முன்னரும் வரும்.

எ.கா.

உகாரம், உங்கள், உச்சி, உடல், உண்டது, உதவி, உந்தி, உப்பு, உமி, உயிர், உரிய, உலகம், உவமை, உளக்கு, உள்ளம், உறவு, உண்ணதம்

உ மொழி இடையில் என்பது *ங்*, *ஞ்* என்ற மெய்கள் தவிர பிற எல்லா மெய்களுக்குப் பின்னரும் வரும்.

எ.கா.

குட்டை, சுடுகாடு, கடுவெளி, கோணும், எய்தும், அந்நுதல், புலி, முகம், யுத்தகாலம், அருள், சொல்லும், அவாவும், அழும், கேளுங்கள், இறுக்கி, அனுகூலம்

உ மொழி இடையில் எல்லா மெய்களுக்கு முன்னரும் வரும்.

எ.கா.

அறுக்க, முங்கு, ஆயுசு, துஞ்சு, இருட்டறை, கருணை, கருதிய, காத்திருத்தேன், கிருபை, கும்பி, குயில், குரங்கு, புல், அரும்பு, முழு, முள்ளு, முறியும், முன்னாள்

உ மொழி இறுதியில் *ங்*, *ஞ்*, *ந்* தவிர பிற எல்லா மெய்களுக்குப் பின்னரும் வரும்.

எ.கா.

அடங்கு, அரசு, ஆடு, கண்ணு, கழட்டு, கருப்பு, உறுமு, தீயு, நற்குரு, நாலு, நாவு, புழு, கொள்ளு, கொன்று, பொன்னு

ஊ

ஊ மொழி முதல், இடை, கடை என்ற மூன்று இடங்களிலும் வரும். *ஊ* மொழி முதலில் *ங்*, *ந்*, *ண்*, *ப்*, *ய்*, *ல்*, *வ்*, *ள்* தவிர எல்லா மெய்களின் முன்னரும் வரும்.

எ.கா.

ஊகார, ஊசி, ஊஞ்சல், ஊட்டாதே, ஊது, ஊமை, ஊர், ஊழி, ஊறு, ஊனம்

ஊ மொழி இடையில் *ங்*, *ஞ்*, *ன்* தவிர பிற எல்லா மெய்களின் பின்னரும் வரும்.

கூகை, சூரர், நிட்டூரன்,தொண்ணூறு, தூங்காமல், நூல், பூசை, மூன்று, மெய்யுயிர், காயலூர், பாலூர், கண்ணூர், கொள்ளுவிர்,

ஊ என்ற உயிர் மொழி இடையில் எல்லா மெய்களுக்கு முன்னரும் வரும்.

எ.கா.

கூகை , தூங்கி, தூசு, பொன்னூஞ்சல், கூட்டம், தூண், கூத்து, பூந்தி, கூப்பிட்டு, பூமி, தூய, தூர, நூல், பூவை,மூழ்கி, தூள், கூனி

எ

எ மொழி முதலிலும் இடையிலும் வரும். *எ* மொழி முதலில் எல்லா மெய்களின் முன்னரும் வரும்.

எ.கா.

எக்காலமும், எங்கள், எச்சம், எஞ்சு, எட்டு, எண்ணம், எதிர், எந்த, எப்போதும், எம், எய்த, எரி, எல்லை, எவர், எழில், எள், எறிகுவோம், என்ன *எ* மொழி இடையில் *ங்*, *ஞ்*, *ன்* தவிர பிற எல்லா மெய்களுக்குப் பின்னரும் வரும். எ.கா.

கெடல், செப்பு, நெட்டெழுத்து, என்னை, தெய்வம், நெஞ்சம், பெண், மெழுகு, வெய்யில்

எ மொழி இடையில் எல்லா மெய்களுக்கும் முன்னரும் வரும்.

எ.கா.

செகம், செங்கதிர், மெச்சு, நெஞ்சு,நெடு, பெண்டிர், மெத்தை, செந்தணல், தெப்பம், தெம்பு, தெய்வம், செருப்பு,செல்லாதே, செவி, அஞ்சழுத்து, தெள்ளு, பெற்றவள், மென்

ஏ

ஏ மொழி முதலில் ஞ் தவிர எல்லா மெய்களின் முன்னரும் வரும்.

எ.கா.

ஏகாந்தம், ஏங்கு, ஏசு, ஏடானை, ஏணி, ஏதுக்கடி, ஏந்து, ஏப்பம், ஏமன், ஏய்க்கும், ஏவல், ஏழு, ஏற்றது

ஏ மொழி இடையில் எல்லா மெய்களின் முன்னரும் வரும்.

எ.கா.

தேக்கு, தேங்கு, தேசம், தேடு, வேண்டு, வேதம், வேந்தன், வேப்பமரம், வேம்பு, வேய், வேர், வேல், வேவு, வேள்வி, வேழம், வேறு, வேனல்

ஏ மொழி இடையில் *ஞ்*, *ன்* தவிர எல்லா மெய்களின் பின்னரும் வரும்.

எ.கா.

கேளாய், அஞ்சேல், கண்டேன், காணேன், தேடல், நேர்மை, பேசும், மேல், செய்யேன், சரேல், வெல்லேன், வேகம், வாழேன், உள்ளேன், வென்றேன்,

ஏ மொழி இறுதியில் *ங்*, *ஞ்*, *ழ்*, தவிர பிற மெய்களின் பின்னர் வரும்.

எ.கா.

ஆவதற்கே, ஆச்சே, இட்டே, பெண்ணே, அடியாதே, பாம்பே, ஆகுமே, அடக்கியே, அறியாதவாரே, அத்தியிலே, அடங்கினவே, கீழே, அடற்குள்ளே, அழுந்திடுவாறே, அவனே

ஐ

ஐ மொழி முதலிலிலும் இடையிலும் கடையிலும் வரும்; மொழி முதலில் *க்*, *ங்*, *ந்*, *ம்*, *ய்*, *வ்* இவற்றின் முன்னர் வரும்.

எ.கா.

ஐக்கியம், ஐங்கரன், ஐந்து, ஐம்பொறி, ஐயர், ஐவர்

ஐ மொழி இடையில் *ங்*, ஞ் தவிர பிற மெய்களின் பின்னர் வரும்.

எ.கா.

ஓகையுடன், சைவம், உடைத்த, துணையதி, தைலம், நைட்டிகர், பையூர், மைப்படும், மாயையடி, விரைந்து, அலைந்து, வைப்பாள், விளையாடு, நனையாதே

ஐ மொழி இடையில் *ங்*, ஞ் தவிர பிற மெய்களின் முன்னர் வரும்.

எ.கா.

வளைக்க, கைசேதம், நைட்டிகர், நினைத்த, நினைந்து, நினைப்பார், பைம்பொன், பொய்மையே, தைலம், நினைவு

ஐ மொழி இறுதியில் ங், ஞ், தவிர எல்லா மெய்களின் பின்னரும் வரும்.

எ.கா.

அங்கை, ஆசை, நிட்டை, கணை, கதை, கிருபை, செந்துரமை, மாயை, முத்திரை, கோவை, ஏழை, காளை, குறை, சேனை

ஒ

ஒ மொழி முதலிலும் இடையிலும் வரும்; மொழி முதலில் ங், ச், ஞ், ந், ம் தவிர பிற மெய்களின் முன்னர் வரும்.

எ.கா.

ஒக்கம், ஒட்டி, ஒண்ணாது, ஒட்டு, ஒப்பனை, ஒயில், ஒரு, ஒலி, ஒவ்வா, ஒழிந்து, ஒளி, ஒறுத்து, ஒன்பது

ஒ மொழி இடையில் *ங்*, ஞ் தவிர பிற மெய்களின் பின்னர் வரும்.

எ.கா.

கொடிமரம், சொல், தொங்கும், நொடி, கண்ணொடு, பொய், மொழி, கையொடு, பேரொளி, தெருவொடு, இதழொடு, கள்ளொடு, அற்றொரு, அவனொடு

ஒ மொழி இடையில் *வ்*, *ர்* தவிர பிற மெய்களின் முன்னர் வரும்.

எ.கா.

பொக்கிசம், பொங்கி, கொச்சை, கொஞ்சு, கொடி, தொண்டு, பொதி, சொந்தம், சொப்பனம், கொம்பு, பொய், பொருள், பொல்லா, தொழும், கொள்ளும், கொன்று

ஓ

ஓ மொழி முதலிலும் இடையிலும் இறுதியிலும் வரும். மொழி முதலில் *க்*, *ங்*, *ச்*, *ட்*, *த்*, *ம்*, *ர்* என்ற மெய்களின் முன்னர் வரும்.

எ.கா.

ஓகை, ஓங்காரம், ஓசை, ஓதில், ஓடு, ஓம், ஓர்

ஓ மொழி இடையில் *ங்*, *ஞ்*, *ண்*, *ர்*, *ள்* தவிர பிற மெய்களின் பின்னர் வரும்.

எ.கா.

கோழி, சோறு, இட்டோம், தோண்டி, நோக்கம், போகம், மோசம், யோகம், நல்லோர், நெய்வோம், உள்ளோர், அற்றோம், தள்ளினோம்

ஓ மொழியிடையில் எல்லா மெய்களுக்குப் முன்னரும் வரும்.

எ.கா.

எமலோகம், போங்காலம், மோசம், உண்டோடி, கோணும், கோதி, போந்தேன், போம், கோயில், பெரியோர், போல, போவான்,கோழி, சோறு, கோன்

ஓ மொழி இறுதியில் *ங்*, *ச்*, *ஞ்*, *ந்*, *ப்*, *ழ்* தவிர பிற எல்லா மெய்களுக்குப் பின்னரும் வரும்.

எ.கா.

எங்கோ, உண்டோ, பாரேணோ, பெரிதோ, போமோ, காணாயோ, செல்வாரோ,

மறந்தல்லோ, பொய்யல்லவோ, போகாளோ, என்றோ, என்னோ?

3.2.4 மெய்யெழுத்துக்களின் வருகை முறை

க்

க் மொழி முதலிலும் இடையிலும் வரும். மொழிமுதலில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

கங்கை, காடு, கிளி, கீழ், குடை, கூகை, கெண்டை, கேணி, கைவிளக்கு, கொம்பு, கோதை, கௌரி

க் மொழி இடையில் எல்லா உயிர்களுக்குப் பின்னரும் வரும்.

எ.கா.

அகம், ஆகாயம், இனிக்குது, சீக்கீரம், உகரம், ஊகாரம், எக்காலமும், ஏகம், ஐக்கியம், ஒக்காளம், ஓகை , ஔகாரம்

க் மொழி இடையில் *ச், ஞ், த், ந், ப், ம், வ்* தவிர எல்லா மெய்களுக்குப் பின்னரும் வரும்.

எ.கா.

இங்கே, கேட்குமடி, வெண்ங்கலம், செய்கை, கார்கோடன், பல்கால், மூழ்கி, கொள்கை, பற்கள், நான்கு

க் மொழி இடையில் எல்லா உயிர்களுக்கு முன்னரும் வரும்.

எ.கா.

தாகம், தாங்காமல், திரிகிறாய், கீரை, சற்குரு, புழுக்கூடு, கெடு, கேடு, மாளிகை, பைகொள், மணக்கோலம்

க்க்

க்க் மொழி இடையில் எல்லா உயிர்களுக்குப் பின்னரும் வரும்.

எ.கா.

தக்கன், தாக்கி, திக்கு, நீக்கல், படுக்கை, மூக்கு, நிக்கு, தேக்கு, தொலைக்க, ஒக்கம், நோக்கம்

க்க் மொழி இடையில் எல்லா உயிர்களுக்கு முன்னரும் வரும்.

எ.கா.

பக்கம், பஞ்சாயாக்காரர், பாக்கியம், தாக்கீது, கொதிக்குது, புழுக்கூடு, தேக்கெடுத்து, வாழ்க்கை, சிவக்கொழுந்து, முக்கோணம்

ங்

ங் மொழி இடையில் மட்டும் வரும்; மொழி இடையில் எல்லா உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

மங்கை, மாங்காய், இங்கே, தீங்கான, துங்க, தூங்காமல், வெங்கதிர், தேங்காய், ஐங்கரன், தொங்கும், போங்காலம்

ங் மொழி இடையில் *க்* முன்னர் மட்டும் வரும்.

எ.கா.

மங்கையர்

ங்ங் மொழி இடையில் வரும்; மொழி இடையில் இ, எ-க்குப் பின்னர் வரும்.

எ.கா.

இங்ஙனம், எங்ஙனம்

ங் மொழி இடையில் அ முன்னர் வரும்.

எ.கா.

எங்ஙனம்

ச் மொழி முதலிலும் இடையிலும் வரும்; மொழி முதலில் எல்லா உயிர்களுக்கு முன்னரும் வரும்.

சடங்கு, சாதி, சிங்கம், சீவன், சுடுகாடு, சூரியன், செவி, சேல்விழி, சைவம், செல், சோறு

ச் மொழி இடையில் ஒ தவிர எல்லா உயிர்களுக்குப் பின்னரும் வரும்.

அசல், ஆசான், இசைந்து, ஈசன், உச்சி, ஊசி, எச்சில், தேசம், கைச்சேதம், ஓசை

ச் மொழி இடையில் எல்லா உயிர்களுக்கு முன்னரும் வரும்.

காஞ்சனை, கஞ்சா, கற்சிலை, திரைச்சீலை, துஞ்சுவதாம், திரிசூலி, அஞ்செழுத்து, ஆதிசேடன், இச்சை, சிற்சொருபம், பரஞ்சோதி, அசௌகரியம்

ச்ச் மொழி இடையில் மட்டும் வரும்; ஐ, ஔ தவிர பிற எல்லா உயிர்களின் பின்னரும் வரும்.

கச்சை, தணாலாச்சு, நிச்சயம், பீச்சு, முச்சந்தி, பூச்சி, மெச்சு, பேச்சு, மொச்சை, போச்சுதே

ச்ச் மொழி இடையில் உயிர்கள் ஊ, ஔ தவிர எல்லா உயிர்களுக்கும் முன்னர் வரும்.

நிச்சயம், பேச்சா, பாய்ச்சி, திரைச்சீலை, பேச்சு, பேச்சே, பேச்செல்லாம், பிச்சை , பேச்சொடு, பேச்சோடு

ஞ்

ஞ் மொழி முதலிலும் இடையிலும் மட்டும் வரும்; மொழி முதலில் ஆ முன்னர் வரும்.

ஞானம்

ஞ் மொழி இடையில் ஏ தவிர எல்லா உயிர்களுக்கும் பின்னர் வரும்.

தஞ்சம், காஞ்சி, பிஞ்சு, ஈஞ்சை, துஞ்சு, ஊஞ்சல், நெஞ்சம், கொஞ்சு, நோஞ்சான்

ஞ் மொழி இடையில் *ஆ* முன்னர் மட்டும் வரும்.

மெஞ்ஞானம்

ஞ் மொழி இடையில் *ச்* முன்னர் மட்டும் வரும்.

அஞ்ச

ஞ்ஞ்

ஞ்ஞ் மொழி இடையில் மட்டும் வரும்; உயிர்கள் *அ*, *இ* இவற்றின் பின்னர் வரும்.

அஞ்ஞானம், விஞ்ஞானம்

ஞ்ஞ் மொழி இடையில் *ஆ* முன்னர் வரும்.

விஞ்ஞானம், மெய்ஞ்ஞானம்

ய்ஞ்ஞ், *ஞ்ச்* மெய் மயக்கத்திலும் வரும்.

மெய்ஞ்ஞானம், அஞ்சு

ட்

ட் மொழி இடையில் மட்டும் வரும்; எல்லா உயிர்கள் பின்னரும் வரும்.

அடி, ஆடு, இடம், ஈட்டி, உடல், ஊடே, எடுத்து, ஏடாய், ஒடி, ஓடு

ட் மொழி இடையில் *ண்* பின்னர் வரும்.

கண்டு

ட் மொழி இடையில் எல்லா உயிர்கள் முன்னரும் வரும்.

கடல் , சடாதரி, சூடியே, காண்டீபன், கூடு, கொடூரன், எட்டெழுத்தாம், கண்டேன், கிடையாதே, வண்டொடு, விட்டோர்

ட் மொழி இடையில் *க்*, *ச்*, *ண்*, *ப்* என்ற மெய்கள் முன்னர் வரும்.

முட்கள், மோட்சம், விட்ணு, நட்பு

ட்ட்

ட்ட் மொழி இடையில் எல்லா உயிர்களுக்குப் பின்னரும் வரும்.

பட்டனம், பெண்டாட்டி, இட்டால் , மீட்டல், முட்டு, மூட்டு, வெட்டு, வேட்டை, நைட்டிகர், பொட்டு, கோட்டை

ட்ட் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்.

அட்டக்கத்தி, அட்டால், விட்டில், விட்டீர், ஆட்டு, நிட்டூரன், நெட்டெழுத்து, விட்டேன், நெட்டை, கட்டொடு, இட்டோம்

ண்

ண் மொழி இடையிலும் இறுதியிலும் வரும். மொழி இடையில் *ஐ* தவிர பிற உயிர்கள் பின்னர் வரும்.

அண்ணாக்கை, ஆண், இணை, வீணாசை, உண்மை, பூண்டு, எண்ணு, ஏணி, ஒண்ணாது, ஓணான்,

ண் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்.

மணல், காணார், திருபணிகள், தண்ணீர், காணும், தொண்ணூறு, எண்ணெய், காணேன், பண்ணொடு, கண்ணோடு, கணை

ண் மொழி இடையில் *க், ட், ப், ம்* முன்னர் வரும்.

வெண்கலம், கண்டான், எண்பது, உண்மை

ண் மொழி இறுதியில் *ஏ, ஐ, ஓ* தவிர்த்து பிற உயிர்களின் பின்னர் வரும்.

நடுவண், காண், விண், வீண், நுண், தூண், பெண்

ண்ண்

ண்ண் மொழி இடையில் மட்டும் வரும்; மொழி இடையில் *அ, இ, உ, எ, ஒ* என்ற உயிர்களின் பின்னர் வரும்.

கண்ணாடி, திண்ணம், உண்ண, எண்ணம், தொண்ணூறு

ண்ண் மொழி இடையில் எல்லா உயிர்கள் முன்னர் வரும்.

வண்ணம், பெண்ணால், புண்ணியம், தண்ணீர், கண்ணு, தொண்ணூறு, எண்ணெய், எண்ணேன், கண்ணை, கண்ணொடு, கண்ணோய்

த்

த் மொழி முதலிலும் இடையில் வரும்; மொழி முதலில் எல்லா உயிர்களுக்கும் முன்னரும் வரும்.

தமிழ், தாய், திருமால், தீ, துறவி, தூசு, தெரு, தேர், தைலம், தொல், தோல்

த் மொழி இடையில் எல்லா உயிர்களின் பின்னரும் வரும்.

அத்தி, ஆத்தாள், இதழ், நன்னீதி, பழுது, ஊது, எத்தனை, ஏதுக்கடி, கைத்தாளம், பொதி, போதகர்

த் மொழி இடையில் *க், ண், ந், ப், ய், ர், ழ், ன்* மெய்களின் பின்னர் வரும்.

முக்தி, எண்திசை, ஏகாந்தம், எய்தும், உதிர்த்து, வாழ்த்தி, உந்தன்

த் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்

கதவு, கூத்தாடி, கொதிக்குது, தடுமாறாதீர், தொழுது,வெகுதூரம், ஆயிரத்தெட்டு, இருந்தேன், ஊத்தை, எண்பத்தொன்று, பெரிதோ

த்த்

த்த் மொழி இடையில் மட்டும் வரும். மொழி இடையில் *ஓ* தவிர பிற உயிர்களின் பின்னர் வரும்.

உற்பத்தி, காத்து, குடித்து, நிறுத்தி, நீத்தார், கூத்து, செத்தார், ஏத்தி, கைத்தாளம்,சொத்துக்கள் .

த்த் மொழி இடையில் *ந்*, *ய்*, *ர்*, *ழ்* என்ற மெய்கள் பின்னர் வரும்.

புகுந்தாள், பொய்த்தாளம், தீர்த்தமும், வாழ்த்தி

த்த் மொழி இடையில் எல்லா உயிர்கள் முன்னர் வரும்.

உய்த்தறி, எடுத்தாளும், கத்தி, காத்தீர், கருத்து, ஆத்தூர், ஆயிரத்தெட்டு, பத்தேடி, ஊத்தை, எண்பத்தொன்று, ஐம்பத்தோர்

ந்

ந் மொழி முதலிலும் இடையிலும் வரும். *ந்* மொழி முதலில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

நஞ்சு, நாகம், நிழல், நீர்க்குமிழி, நுண், நூல், நெற்றி, நேர்மை, நைட்டிகர், நொடி, நோக்கம்

ந் மொழி இடையில் எல்லா உயிர்களின் பின்னரும் வரும்.

எ.கா.

அந்த, ஆந்தை, இந்த, நீந்து, கொழுந்தி, கூந்தல், செந்தணல், வேந்தன், ஈரைந்து, சொந்தம், போந்தேன்

ந் மொழி இடையில் *ய்*, *ர்*, *ழ்* பின்னர் வரும்.

எ.கா.

மேய்ந்து, அமர்ந்து, மகிழ்ந்து

ந் மொழி இடையில் *ஐ* தவிர்த்த பிற உயிர்களின் முன்னர் வரும்.

எ.கா.

நன்கு, பதிநான்கு, தவநிலை, முந்நீர், அந்நுதல், அந்நூல், அந்நெஞ்சு, அந்நொடி, அந்நோக்கம்,

ந் மொழி இடையில் *த்* முன்னர் மட்டும் வரும்.

எ.கா.

படர்ந்து

ந்ந் மொழி இடையில் மட்டும் வரும். மொழி இடையில் உ, எ உயிர்கள் பின்னர் வரும்.

எ.கா.

முந்நீர், செந்நீர்

ந்ந் மொழி இடையில் ஆ, ஈ உயிர்கள் முன்னர் வரும்.

எ.கா.

எந்நாளும், செந்நீர்

ப்

ப் மொழி முதலிலும் இடையிலும் வரும்; மொழி முதலில் எல்லா உயிர்களுக்கும் முன்னர் வரும்.

எ.கா.

பகல், பாடல், பிணி, பீற்றல், புலி, பூட்டு, பெரு, பேய், பை, பொய், போதகர், பௌர்ணமி

ப் மொழி இடையில் எல்லா உயிர்களுக்கும் மெய்களுக்கும் பின்னர் வரும்.

எ.கா.

அப்பு, சாபம், பறிப்பான், தீப்பொறி, அறுபத்துநாலு, கூப்பிடு, செப்பு, ஏப்பம், களைப்பு, ஒப்பம், கோபுரம்

ப் மொழி இடையில் ட், ண், ம், ய், ர், ல், ழ், ள், ன் என்ற மெய்களுக்குப் பின்னர் வரும்.

எ.கா.

கட்புலன், பண்புடனே, பாம்பு, பொய்மை, மார்பு, தொல்பொருள், வாழ்பவர், நீள்பதம், அன்பு

ப் மொழி இடையில் எல்லா உயிர்களுக்கும் முன்னர் வரும்.

எ.கா.

அன்பன், கொடுப்பான், செபித்து, திண்பீராகில், பிறப்பு, பெரும்பூதம், பேதைப்பெண்ணாம், அகப்பேய், பை, ஐப்பொறி, படிப்போர்

ப்ப் மொழி இடையில் மட்டும் வரும்; மொழி இடையில் எல்லா உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

இடப்பாகன், காஞ்சப்புகை, கழிப்பார், தீப்பொறி, முப்பால், கூப்பிடு, செப்பு, ஏப்பம், நினைப்பார், ஒப்பனை, தோப்பிலே

ப்ப் மொழி இடையில் *ர், ய், ழ்* என்பவைகளுக்குப் பின்னர் வரும்.

எ.கா.

சேர்ப்போம், மெய்ப்பொருள், காழ்ப்பு

ப்ப் மொழி இடையில் எல்லா உயிர்களுக்கும் முன்னர் வரும்.

எ.கா.

எழுப்ப, இருப்பார், பரப்பி, வைப்பீரே, அப்பு, அப்பூ, பேதை, அகப்பேய்

தீப்பொறி, அடைப்போம், தொப்பை

ம்

ம் மொழி முதலிலிலும் இடையிலும் கடையிலும் வரும். மொழி முதலில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

மக்கள், மான், மின்னர், மீன், முல்லை, மூன்று, மெய்ஞானம், மேல், மை, மொழி, மோனம்

ம் மொழி இடையில் எல்லா உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

அம்பு, ஆம்பல், நிமிர், வீம்பு, அரும்பு, ஊமை, எமன், ஏமன், நிலைமை, கொம்பு, கோம்பை

ம் மொழி இடையில் *ண், ய், ர், ன்* மெய்களின் பின்னர் வரும்.

எ.கா.

உண்மை, வாய்மை, நேர்மை, நன்மை

ம் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

அமர், அம்மா, அம்மி, தம்மீது, அமுதம், அம்மூவர், நம்மெதிரே, தாமே, நம்மை, அம்மொழி, தம்மோடு

ம் மொழி இடையில் *ப்* மெய்யின் முன்னர் வரும்.

எ.கா.

வேம்பு

ம் மொழி இறுதியில் *இ*, *ஈ*, *ஊ*, *ஏ*, *ஐ*, *ஓ* தவிர பிற எல்லா உயிர்களின் பின்னரும் வரும்.

எ.கா.

குற்றம், தாம், அமரும், எம், எரிப்போம்

ம்ம் மொழி இடையில் மட்டும் வரும். *ம்ம்* மொழி இடையில் *ஈ*, *ஊ*, *ஏ*, *ஓ* உயிர்கள் தவிர பிற உயிர்களுக்குப் பின்னர் வரும்.

எ.கா.

தம்மை, தாமென்று, இம்மென்று, உம்மென்று, செம்மறி, பொம்மை,

ம்ம் மொழி இடையில் எல்லா உயிர்களுக்கும் முன்னரும் வரும்.

எ.கா.

மும்மலம், சும்மா, கும்மி, தம்மீது, தம்முடைய, அம்மூன்று, இம்மென்று, தம்மேனி, தம்மொடு, தம்மோடு, தம்மை

ய்

ய் மொழி முதலிலும் இடையிலும் கடையிலும் வரும்; *ய்* மொழி முதலில் *அ*, *ஆ*, *உ*, *ஊ*, *ஓ* என்ற உயிர்களுக்கு முன்னர் வரும்.

எ.கா.

யமபுரம், யானை, யுத்தகாலம், யூகம், யோகம்

ய் மொழி இடையில் எல்லா உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

வயல், வெண்காயம், வேதியர், ஈயம், குயில், தூய, தெய்வம், நேயம், தையல், ஒயில், கோயில்

ய் மொழி இடையில் எல்லா உயிர்களுக்கும் முன்னர் வரும்.

எ.கா.

ஆகாயம், எய்யாமல், ஒயில், செய்யீர், செய்யும், பையூர், மெய்யென்று, அறியேன், கிரியை, மெய்யொழுக்கம், பெரியோர்

ய் மொழி இடையில் *க்*, *ச்*, *த்*, *ந்*, *ப்*, *ம்*, *வ்* என்ற மெய்களுக்கு முன்னர்.

எ.கா.

செய்கை, பாய்ச்சி, எய்தல், மேய்ந்து, பொய்ப்பொருள், பாய்மரம், செய்வான்

ய் மொழி இறுதியில் எல்லா உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

கய், வாய், பிய், தீய், புயல், துய், மெய், பேய், பையன், பொய், ஓய்

ய்ய் மொழி இடையில் மட்டும் வரும். *ய்ய்* மொழி இடையில் *எ*, *உ*, *ஒ* என்ற உயிர்களுக்குப் பின்னர் வரும்.

எ.கா.

எய்யாமல், துய்ய, பொய்யான

ய்ய் மொழி இடையில் எல்லா உயிர்களுக்கு முன்னரும் வரும்.

எ.கா.

மெய்யடி, மெய்யாக, நெய்யில், செய்யீர், செய்யும், மெய்யூர், மெய்யென, மெய்யே, மெய்யை, மெய்யொடு, நெய்யோடு.

ர்

ர் மொழி இடையிலும் இறுதியிலும் மட்டும் வரும். மொழி இடையில் எல்லா உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

அரிசி, ஆராய்ந்து, இரவு, இராறு, உருவம், ஊர், எரி, ஏரி, ஒருமை, ஓர், ஐராதம்

ர் மொழி இடையில் எல்லா உயிர்களுக்கு முன்னரும் வரும்.

எ.கா.

கரம், சாராமல், சூரியன், வாரீர், புருவம், கரூர், யாரென, யாரேனும், விரைந்து, பேரொளி, பாரோர்

ர் மொழி இடையில் *க் ச்*, *த்*, *ந்*, *ப்*, *ம்*, *வ்* என்ற மெய்களுக்கு முன்னர் வரும்.

குதர்க்கம், மூர்ச்சை, சேர்த்து, கூர்ந்து, சேர்ப்போம், சேர்மின், சேர்வோர்

ர் மொழி இறுதியில் ஐ தவிர பிற எல்லா உயிர்களுக்குப் பின்னரும் வரும்.

எ.கா.

தேசிகர், நகார், உயிர், எண்ணுவீர், உரு, ஊர், நெருப்பு, தேர், அவனொரு, அண்டினோர்

ல்

ல் மொழி இடையிலும் இறுதியிலும் மட்டும் வரும்; மொழி இடையில் எல்லா உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

சலனம், காலம், சிலை, நீலம், துலாவி, நூல், நெல், வேலை, தைலம், தொல்லை, போல

ல் மொழி இடையில் தவிர எல்லா உயிர்களுக்கும் முன்னர் வரும்.

எ.கா.

அம்பலம், அல்லால், அம்புலி, கல்லீரல், பாலு, காயலூர், மெல்லென, கல்லேட்டில், அலைந்து, கல்லொடு, எல்லோரும்

ல் மொழி இறுதியில் *அ*, *ஆ*, *இ*, *ஏ* என்ற உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

வாசல், மால், எழில், மேல்

ல்ல் மொழி இடையில் மட்டும் வரும்; மொழி இடையில் அ, இ, உ, ஒ என்ற உயிர்களுக்குப் பின்னர் வரும்.

எ.கா.

அல்லல், இல்லை, முல்லை, எல்லை, கொல்லன்

ல்ல் மொழி இடையில் எல்லா உயிர்களுக்கும் முன்னர் வரும்.

எ.கா.

அல்லல், அல்லால், மெல்லி, சொல்லீர், கொல்லு, கல்லூர், கல்லென, கல்லேன், முல்லை, கல்லொடு, எல்லோரும்

வ்

வ் மொழி முதலிலும் இடையிலும் வரும்; மொழி முதலில் *உ*, *ஊ*, *ஒ*, *ஓ* தவிர பிற எல்லா உயிர்களின் முன்னர் வரும்.

எ.கா.

வயசு, வாசம், விதி, வீடு, வெயில், வேம்பு, வையம்

வ் மொழி இடையில் எல்லா உயிர்களுக்கும் பின்னர் வரும்.

எ.கா.

அவள்,ஆவி, இவன், சிவன், சீவல், உவமை, தூவி, எவர், ஏவல்

வ் மொழி இடையில் *ய், ர், ல், ழ், ள், ஸ்* பின்னர் வரும்.

எ.கா.

தெய்வம், சார்வார், செல்வார், தாழ்வை, கேள்வி, சரஸ்வதி

வ் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

சிவம், சாவார், அவித்த, தேய்வீகம், சாவு, பாவூர், வெட்டவெளி, வெவ்வேறு, வேண்டாவே, வேண்டாவோ

வ்வ்

வ்வ் மொழி இடையில் மட்டும் வரும். *வ்வ்* மொழி இடையில் *அ, எ, இ, ஒ* என்ற உயிர்களின் பின்னர் வரும்.

எ.கா.

அவ்வை, வெவ்வினை, இவ்வண்ணம், ஒவ்வாமை

வ்வ் மொழி இடையில் *அ, ஆ, இ, உ, ஒ* என்ற உயிர்களுக்கு முன்னர் வரும்.

எ.கா.

எவ்வகையாக, ஒவ்வாமை, வெவ்வினை, எவ்வுலகும், ஒவ்வொரு

ழ்

ழ் மொழி இடையிலும் இறுதியிலும் மட்டும் வரும்; ஐ தவிர பிற எல்லா உயிர்களின் பின்னரும் வரும்.

எ.கா.

அழித்து, ஆழ்ந்து, இழந்து, கீழாக, புழு, ஊழ்வினை, தொழுது, கோழி

மொழி இடையில் எல்லா உயிர்கள் முன்னரும் வரும்.

எ.கா.

சூழ, சூழா, வேல்விழி, வாழீர், தாழும், பாழூர், இதழென, வாழேனோ, ஏழை

ழ் மொழி இடையில் *க், த், ந், ப், வ்* முன்னர் வரும்.

எ.கா.

வாழ்க , வாழ்த்தி, வாழ்ந்த, வாழ்பவர், வாழ்வு

ழ் மொழி இறுடதயில் *உ, எ, ஏ, ஐ, ஒ, ஓ* தவிர பிற எல்லா உயிர்கள் பின்னர் வரும்.

எ.கா.

இதழ், பாழ், முத்தமிழ், கீழ், சூழ்

ள்

ள் மொழி இடையிலும் இறுதியிலும் வரும்; மொழி இடையில் *ஐ* தவிர பிற உயிர்களுக்குப் பின்னர் வரும்.

எ.கா.

அளவு, ஆளலாம், இளப்பம், நீள்படம், குளவி, தூளால், தெளிந்து, கேள்வி, கொள்கை, கோள்கள்

ள் மொழி இடையில் எல்லா உயிர்களுக்கும் முன்னர் வரும்.

எ.கா.

மிளகு, மீளாதே, ஒளி, வெள்ளீயம், கொள்ளும், உள்ளூர், உள்ளே, களைப்பு, உள்ளோர்

ள் மொழி இடையில் *க்*, *ப்*, *வ்* என்ற மெய்களின் முன்னர் வரும்.

எ.கா.

கொள்கை, நீள்படம், கேள்வி

ள் என்ற மெய் மொழி இறுதியில் *இ*, *ஈ*, *ஐ* தவிர பிற உயிர்களின் பின்னர் வரும்.

எ.கா.

எங்கள், எரித்தாள், கடவுள், தூள், எள், தேள், கொள், கோள்

ள்ள் மொழி இடையில் மட்டும் வரும்; *ள்ள்* மொழி இடையில் *அ*, *இ*, *உ*, *எ*, *ஒ* என்ற உயிர்களின் பின்னர் வரும்.

எ.கா.

கள்ள, பிள்ளை, உள்ள, வெள்ளி, கொள்ள

ள்ள் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

கள்ளம், தள்ளாதே, வெள்ளி, உள்ளீர், கொள்ளு, உள்ளூர், கள்ளென, உள்ளே, வெள்ளை, கள்ளொடு, உள்ளோர்

ற்

மொழி இடையில் மட்டும் வரும். *ற்* மொழி இடையில் *ஐ* தவிர பிற உயிர்களின் பின்னர் வரும்.

எ.கா.

அற்புதம், ஆறு, இறப்பு, கீறி, குற்றம், கூறு, நெறி, வேறு, ஐம்பொறி, பொறு, சோறு,

ற் மொழி இடையில் எல்லா உயிர்களின் முன்னர் வரும்.

எ.கா.

அற்றவர், ஆறாறும், ஊறிடும், ஈறாறு, கற்றூண் நூற்றேட்டும், இன்றே, மறைத்து, என்றொரு, பெற்றோர்,

ற் மொழி இடையில் *ன்* என்ற மெய்யின் பின்னர் வரும்.

எ.கா.

குன்று

ற் என்ற மெய் மொழி இடையில் *க்*, *ச்*, *ப்* என்ற மெய்களின் முன்னர் வரும்.

எ.கா.

இயற்கை, நாற்சதுரம், பொற்பதம்

ற்ற்

ற்ற் மொழி இடையில் மட்டும் வரும். மொழி இடையில் *ஐ* தவிர பிற உயிர்களின் பின்னர் வரும்.

எ.கா.

அகற்றி, ஆற்றும், சிற்றடி, பீற்றல், முற்றும்,ஊற்றி, நெற்றி, நேற்றி, கொற்றவன், தோற்றம்

ற்ற் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

நாற்றம், பற்றாது, பற்றி, போற்றீரே, அற்று, வேற்றூர், நூற்றெட்டும், குற்றேவல், ஊற்றை, மற்றொரு, பற்றற்றோன்

ன்

ன் மொழி இடையிலும் இறுதியிலும் வரும். மொழி இடையில் *ஐ* தவிர பிற உயிர்களின் பின்னர் வரும்.

எ.கா.

அன்னை, ஆனால், இனி, ஈன்று, உனக்கு, ஊனம், என்ன, ஏனென்று, கொன்றை, கோனார்

ன் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

சயனம், அன்னார், சூனியம், நன்னீதி, பொன்னு, கன்னனூர், நன்னெறி, மானே, யானை, என்னொடு, என்னோடு

ன் மொழி இடையில் *க், த், ப், ம், ய், ற்* என்ற மெய்களின் முன்னர் வரும்.

எ.கா.

நான்கு, அவன்தான், அன்பு, ஆன்மா, தன்யம், அகன்று

ன் மொழி இறுதியில் எல்லா உயிர்களின் பின்னர் வரும்.

எ.கா.

ஏகம், கடந்தான், நின், மீன், உன், காற்றூண், என், சொல்கிறேன், செம்பொன், வந்தோன்

ன்ன் மொழி இடையில் *ஆ, ஈ, ஊ, ஏ, ஐ, ஓ* என்ற உயிர்கள் தவிர பிற உயிர்களின் பின்னர் வரும்.

எ.கா.

அன்னம், இன்னம், உன்னதம், என்னோ, சொன்ன

ன்ன் மொழி இடையில் எல்லா உயிர்களின் முன்னரும் வரும்.

எ.கா.

அன்னம், சொன்னார், பன்னிரண்டு, நன்னீதி, பொன்னு, பொன்னூர், நன்னெறி, சொன்னேன், அன்னை, என்னொடு, என்னோடு

3.2.5 மெய்மயக்கம் (Consonant Cluster)

சொற்களில் இரு மெய்கள் தம்முள் இணைந்து வரும்போது அதனை மெய்மயக்கம் என்கின்றோம். உடனிலை மயக்கம் (ஒத்த மெய்கள் இணைந்து வருவது), வேற்றுநிலை மயக்கம் (தம்முள் வேறு வேறு மெய்கள் இணைவது) என மெய்மயக்கம் இரண்டு வகைப்படும். மொழியியல், இவ்வாறு மொழிகளுக்கிடையில் வரும் மெய் ஒலிகள் இணைந்து வரும் நிலையை நான்காக வகைப்படுத்தும். தமிழில் இரு மெய்களின்

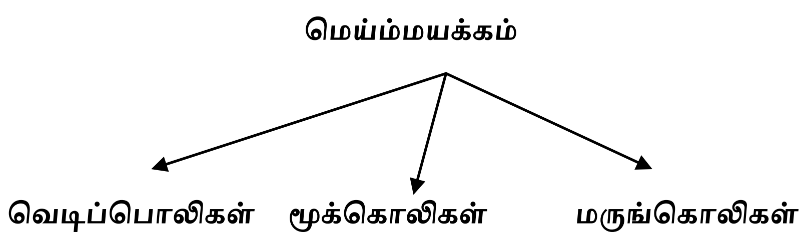
மயக்கமும் மூன்று மெய்களின் மயக்கமும் உள்ளன. மெய் மயக்கங்கள் மொழி இடையில் மட்டும் வரும்.

(I) இரு மெய்களின் மயக்கம்

1. ஒத்த இரு மெய்கள் உடனிலை மயக்கம் C_1 C_1 (Identical cluster)
2. வேறுபட்ட இரு மெய்கள் வேற்றுநிலை மயக்கம் C_1 C_2 (Non Identical Cluster)

(II) மூன்று மெய்களின் மயக்கம்

3. வேறுபட்ட ஒரு மெய்யும் ஒத்த இரு மெய்யும் C_1 C_2 C_2
4. மூன்று வேறுபட்ட மெய்கள் C_1 C_2 C_3



ஒத்த இரு மெய்கள் இணைந்து வரும் மெய்மயகங்களை மூன்றாகப் பகுக்கலாம் வெடிப்பொலிகளின் மயக்கம், மூக்கொலிகளின், மருங்கொலிகளின் மயக்கம் ஆகியவை ஆகும். அவற்றைப் பின் வருமாறு பட்டிய்லிடலாம்.

வெடிப்பொலிகள் (Plosives sounds)

க்க்	அக்காள்
ச்ச்	அச்சு
ட்ட்	ஈட்டி
த்த்	அத்தி (யானை)
ப்ப்	இடப்பாகன்
ற்ற்	பற்றாது

மூக்கொலிகள் Nasal sounds

ங்ங்	இங்ஙன்
ஞ்ஞ்	அஞ்ஞானம்
ண்ண்	எண்ணம்
ந்ந்	செந்நீர்
ம்ம்	சும்மா
ன்ன்	அன்னை

மருங்கொலிகள் (Lateral Sounds)

ல்ல்	இல்லறம்
ள்ள்	கள்ளர்
ய்ய்	பொய்யாது
வ்வ்	அவ்வை

உடனிலை மெய்களின் வருகை அட்டவணை C_1C_1

	க்	ங்	ச்	ஞ்	ட்	ண்	த்	ந்	ப்	ம்	ய்	ர்	ல்	வ்	ழ்	ள்	ற்	ன்
க்	X																	
ங்		X																
ச்			X															
ஞ்				X														
ட்					X													
ண்						X												
த்							X											
ந்								X										
ப்									X									
ம்										X								
ய்											X							

ர்												--						
ல்													X					
வ்														X				
ழ்															--			
ள்																X		
ற்																	X	
ன்																		X

வேற்றுநிலை மெய்ம்மயக்கம் (இரு வேறுபட்ட மெய்களின் வருகை) C_1C_2

இவ்வமைப்பில் இரு வேறுபட்ட மெய்களின் வருகைகள் வெடிப்பொலி, மூக்கினவொலி, மருங்கொலி, அரையுயிர், வருடொலி ஆகியவற்றுடன் இடம்பெற்றுள்ளன. அவை கீழே வகைப்படுத்தித் தரப்பட்டுள்ளன.

க்த்	பக்தி
ட்க்	கேட்கும்
ட்ச்	மோட்சம்
ட்ப்	நட்பு
ற்க்	இயற்கை
ற்ப்	நிற்போம்
ட்ண்	விட்ணு

ங்க்	அங்கை
ஞ்ச்	அஞ்சி
ண்க்	வெண்கலம்
ண்ட்	ஆண்டவன்
ண்ப்	உண்போம்
ண்ம்	அரண்மனை
ந்த்	ஆந்தை

ம்ப்	இரும்பு
ன்க்	நான்கு
ன்ப்	அன்பர்
ன்ம்	சன்மார்க்கம்
ன்ற்	வென்றான்

ல்ப்	இயல்பு
ல்வ்	செல்வி
ள்க்	கொள்கை
ள்ப்	நீள்பதம்
ள்வ்	கேள்வி
ழ்க்	வாழ்க
ழ்ப்	வாழ்பவர்
ழ்வ்	வாழ்வோம்
ழ்ந்	அமிழ்ந்து
ய்க்	ஏய்க்கும்
ய்த்	செய்தேன்
ய்ம்	பொய்மை
ய்வ்	செய்வான்
ர்க்	கார்கோடகன்
ர்த்	தேனாமிர்தம்
ர்ப்	மார்பு
ர்ம்	நேர்மை
ர்வ்	பார்வதி

மெய்+க்	ங்க் , ட்க், ணக், ய்க், ர்க், ழ்க், ள்க், ற்க், ன்க்
மெய்+ச்	ஞ்ச், ட்ச்
மெய்+ட்	ண்ட்
மெய்+ண்	ட்ண்
மெய்+த்	க்த், ந்த்,ய்த்,ர்த்,

மெய்+ந்	ழ்ந்
மெய்+ப்	ட்ப், ண்ப், ம்ப், ர்ப், ல்ப், ழ்ப், ள்ப், ற்ப், ன்ப்.
மெய்+ற்	ன்ற்
மெய்+ம்	ண்ம், ய்ம், ர்ம், ன்ம்
மெய்+வ்	ய்வ், ர்வ், ல்வ், ழ்வ், ள்வ்

வேற்றுநிலை இருமை மெய்மயக்கங்கள்

இரண்டாவது உறுப்பினர்

முதலாவது உறுப்பினர்

	க்	ச்	ட்	த்	ப்	ற்	ங்	ஞ்	ண்	ந்	ம்	ன்	ய்	ர்	ல்	வ்	ழ்	ள்	ஸ்
க்				X															
ச்																			
ட்	X	X			X				X										
த்																			
ப்																			
ற்	X				X														
ங்	X																		
ஞ்		X																	
ண்	X		X		X						X								
ந்				X															
ம்					X														
ன்	X				X	X					X								
ய்	X			X							X					X			
ர்	X			X	X						X					X			

ல்					X											X			
வ்																			
ழ்	X				X					X						X			
ள்	X				X											X			

மூன்று மெய்யெழுத்து மயக்கம் (Three Consonant Clusters $C_1 C_2 C_2$)

வேறுபட்ட மெய்யுடன் ஒத்த இரு மெய்கள் இணையும் வடிவங்கள் $C_1 C_2 C_2$

வாழ்க்கை
வாழ்த்தி
சர்க்கரை
சேர்த்து
பார்ப்போம்
உணர்ச்சி
ஏய்க்கும்
வாய்த்து
பாய்ச்சி

மூன்று வேறுபட்ட மெய்களின் இணைவு வடிவம் $C_1 C_2 C_3$

இவ்வடிவம் வருடொலி, அரையுயிர், மருங்கொலி ஆகியவற்றில் இணைந்து வருகின்றன.

ஆய்ந்து
ஓர்ந்து
சூழ்ந்திடும்

3.2:6 அசை அமைப்பு (Syllabic Structure)

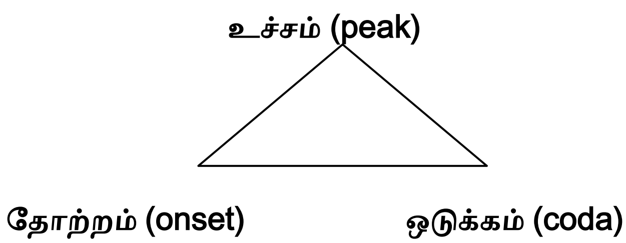
அசை அமைப்பு ஒலியனியல் அமைப்பின் மற்றொரு பகுதியாகும். அசை அமைப்பு உயிர்கள் அடிப்படையிலானது. அசையில் முன்னலகு, இடையலகு, கடையலகு என்ற மூன்று பகுதிகள் உள்ளன. அசையின் இடையலகுதான் கருப்பகுதி. ஒரு சொல்லில் ஒன்றொ அதற்கு மேற்பட்ட அலகுகள் வரலாம். ஒவ்வொரு மொழியும் அசை அமைப்பில் வேறுபடும். தமிழுக்கென்று தனிப்பட்ட அசை அமைப்பு உள்ளது. ஒன்பது வகையான அசை அமைப்புகள் உள்ன: **V̄, V̄C, CVC, CV̄, CV̄C, V, VC, CV, CVCC, C V̄CC**

மொழிக்கு அடிப்படையாக அமைவனவற்றுள் அசையும் ஒன்று. சொற்களில் காணப்படும் உயிர் ஒலிகளை மையமாகக் கொண்டு அதற்கு முன்னும் பின்னும் மெய்யெழுத்துக்கள் வந்தும் அல்லது வரமாலும் இருக்கும் நிலையில் அசை உருவாக்கம் பெறுகிறது.

அசை (Syllable)

ஒவ்வொரு உயிரும் ஓர் அசையாகும். ஒரு சொல்லில் ஒன்றுக்கு மேற்பட்ட எண்ணிக்கையில் அசைகள் அமையலாம். ஒலியசைக்குத் தோற்றம் உச்சம் ஒடுக்கம் என்ற மூன்று உறுப்புக்கள் உண்டு.

C_1 C_2 C_3



உச்சம் கட்டாய உருப்பாகும். ஏனைய இரண்டும் விருப்ப உறுப்புகளாகும். உயிர் ஒலியன் உச்ச உறுப்பாகும், மெய் ஒலியன் (consonant phoneme) தோற்றம், ஒடுக்க உறுப்பாக வரும். ஆய்வுக்கு உட்பட்ட பகுதியில் ஓரசை முதலாக ஒன்பதசை வரையிலான சொற்கள் பிரித்து வகைப்படுத்தப்பட்டுள்ளன.

வ.எ	அசை அமைப்பு	சொல்
1	CVC	நல்
2	CV̏	பூ
3	CV̏C	காண்
4	VC	எண்
5	V̏C	ஊர்
6	V̏	ஈ

ஈரசைச் சொற்கள் 35 அமைப்பு வகை

வ.எ S:N	அசை அமைப்பு	சொல் சொல் word
1.	CVC,CV	கல்வி
2.	CVC,CVC	தெய்வம்
3.	CV̏CC,CV	வாழ்த்தி
4.	CV̏,CVC	காதல்
5.	CV̏,CV̏	கோதை
6.	VC,CVC	அன்பர்
7.	V,CVC	அகம்
8.	CV̏C,CV̏C	வாழ்வேன்
9.	CVC,CV̏	தும்பீ
10.	CVC,CV̏C	சுக்கான்
11.	CV,CV	புலி

12.	CV,CVC	தமிழ்
13.	CV,CV̏	திசை
14.	VC,CV	உண்டு
15.	V,CV̏	இறை
16.	V̏,CVC	ஈசன்
17.	V̏C,CV	ஐந்து
18.	V,CV	அணி
19.	V̏,CV̏	ஆசை
20.	V̏,CV̏C	ஆசான்
21.	CV̏C,CV̏	நேர்மை
22.	CV̏CC,CV̏C	சேர்ப்போம்
23.	CV̏CC,CVC	மார்க்கம்
24.	V̏C,CV̏	ஆந்தை
25.	V̏C,CVC	ஆட்டம்
26.	V̏CC,CV	ஆய்ந்து
27.	V̏C,CV̏C	ஆத்தாள்
28.	CV̏CC,CV̏	வாழ்க்கை
29.	CV̏C,CV	பாம்பு
30.	CV̏,CV	சோதி
31.	VC,CV̏	உண்மை
32.	VC,CV̏C	உண்பாய்
33.	V̏,CV	ஈது
34.	CV̏C,CVC	கூந்தல்
35.	CV̏,CV̏C	பையூர்

மூவசைச் சொற்கள்

வ.எ	அசை அமைப்பு	சொல்
1.	CV̏C,CV̏,CVC	தாங்காமல்
2.	CVC,CV,CV	செம்மறி
3.	CV̏,CV,CVC	பூரணம்
4.	CVC,CV,CVC	சின்மயம்
5.	CV̏,CV̏,CV	வையாளி
6.	CV̏,CV̏,CVC	பூலோகம்
7.	CV̏C,CV,CVC	சீக்கிரம்
8.	CVC,CV̏,CV	சிக்காது
9.	CVC,CV̏,CVC	வெங்காயம்
10.	CV,CVC,CV	மருந்து
11.	CV,CV,CV	துறவி
12.	CV,CVC,CV̏C	குதம்பாய்
13.	CV,CV,CVC	கடவுள்
14.	CV,CVC,CVC	விளங்கும்
15.	CV̏C,CV,CV	பார்வதி
16.	CVC,CV,CV̏C	சித்தனார்
17.	CV,CV,CV̏	மகிமை
18.	V,CVC,CV	அறிந்து
19.	CV,CV,CV̏C	பெரியோர்
20.	CV,CV̏C,CV	விரைந்து
21.	CV,CV̏,CV	நினைவு
22.	CV,CV̏,CVC	தலைவன்
23.	V,CV,CV	அரிசி

24.	CV̏,CV,CV̏	தாமரை
25.	CV̏,CV,CV̏C	ஞானியார்
26.	CV̏,CV̏,CV̏	காணாதே
27.	VC,CV,CVC	அங்குசம்
28.	VC,CV̏,CV	அஞ்சாத
29.	VC,CV̏,CVC	உண்ணாமல்
30.	V,CVC,CV̏C	அகப்பேய்
31.	V,CV,CVC	உலகம்
32.	V,CV,CV̏C	அறிவாய்
33.	V̏,CV,CVC	ஆணவம்
34.	V̏,CV̏,CVC	ஆகாயம்
35.	V,CV̏C,CV̏C	உரைப்பார்
36.	V,CVC,CVC	அமிழ்தம்
37.	VC,CV̏,CV̏	ஒத்தாசை
38.	VC,CV CV	உத்தமி
39.	CV̏,CVC,CVC	மாணிக்கம்
40.	CVC,CV,CV̏	வல்லமை
41.	CVC,CV̏C,CV	பல்லாக்கு
42.	CVC,CV̏,CV̏	விள்ளாதே
43.	V̏,CV,CV	ஓதிய
44.	V̏,CV,CV̏C	ஓடுவார்
45.	V,CV̏C,CV	இசைந்து
46.	CV̏C,CV,CV̏C	பூட்டுவோம்
47.	CV̏C,CV̏,CV	நீங்காது
48.	CV̏C,CV̏,CV̏	காட்டாதே
49.	CV̏,CVC,CV	பூசித்து

50.	CV̏CC,CV,CV	பார்ப்பது
51.	CV̏CC,CV,CVC	சூழ்ந்திடும்
52.	V̏C,CV,CV	ஐம்பொறி
53.	CVC,CVC,CV	கண்டித்த
54.	CVC,CVC,CV̏C	கண்கண்டேன்
55.	V,CV̏,CV	உபாதி
56.	V,CV̏,CVC	இறைவன்
57.	V,CVCC,CV	உணர்ந்து
58.	V,CV,CV̏	அருமை
59.	VC,CV,CV̏	அம்பிகை
60.	VC,CV̏C,CV	அண்ணாக்கு
61.	V̏,CV̏,CV	ஈராறு
62.	V,CV̏,CV̏C	உறைவார்
63.	V,CVC,CV̏	இயற்கை
64.	VC,CV̏,CV̏C	இல்லையேல்
65.	CV̏,CV̏C,CV	வீறாப்பு
66.	CV,CV̏C,CV̏C	நினைப்பார்
67.	CV,CVCC,CV	புகழ்ந்து
68.	CVC,CVC,CVC	தித்திக்கும்
69.	CVC,CV̏C,CVC	சித்தாந்தம்
70.	V̏C,CV,CV̏C	ஆக்குவோம்
71.	V̏,CV̏,CV̏	ஓடாதே
72.	V̏C,CV,CVC	ஐங்கரன்
73.	V̏CC,CV,CV	ஏய்க்குது
74.	CV̏CC,CVC,CV̏C	தீர்ந்தக்கால்
75.	CV̏C,CV,CV̏	நான்மறை

76.	CV̏,CV̏C,CVC	பாராட்டும்
77.	CV,CVC,CV̏	படுக்கை
78.	CVC,CVC,CV̏	கண்ணம்மா
79.	CVC,CV̏CC,CVC	சன்மார்க்கம்
80.	CVC,CV̏C,CV̏C	செய்யாக்கால்
81.	CV,CVCC,CV̏C	கவிழ்ந்தேன்
82.	CV̏C,CVC,CV	நாற்பத்து
83.	VC,CVC,CV	உற்பத்தி
84.	VC,CVC,CV̏C	உட்கொண்டேன்
85.	VC,CV,CV̏C	எண்ணுவீர்
86.	V̏C,CV̏,CV	ஆத்தாடி
87.	V̏C,CV̏,CVC	ஓங்காரம்
88.	V̏,CVC,CVC	ஆனந்தம்
89.	V̏,CV̏C,CV	ஈரைந்து
90.	V̏,CV̏C,CVC	ஏகாந்தம்
91.	V̏C,CV̏,CV̏	ஊட்டாதே
92.	CV̏,CV̏,CV̏C	சாராதார்
93.	CV̏, CV, CV	காதலி

ஐந்தசை சொற்கள்

வ.எ	அசை அமைப்பு	சொல்
	CV,CVC,CV,CVC,CV	தவிக்கிறண்டி
	CV,CVC,CV,CV̏,CVC	கடக்கலாகும்
	CV,CVC,CV,CV,CV̏C	தடுத்திடலாம்

	V,CVC,CV,CV,CV	அகற்றிவிடு
	V,CVC,CV,CVC,CV	அடுத்தடுத்து
	CV,CVC,CV,CVC,CV̏C	மறந்திருப்போம்
	CV,CV,CV,CV,CV	கடுகளவு
	CV,CV,CV̏C,CV,CVC	பிரமாந்திரம்
	VC,CV,CV,CV̏,CVC	இந்திரசாலம்
	V,CVC,CV,CV,CV̏C	அறிந்திடவாம்
	V,CVC,CVC,CV,CV	அகப்பட்டது

ஆறசைச் சொற்கள்

வ.எ S:N	அசை அமைப்பு structure of syllabic	சொல் word
	CV,CVC,CV,CV,CV̏,CVC	பகுத்தறியாமல்

ஒன்பதசைச் சொற்கள்

வ.எ S:N	அசை அமைப்பு structure of syllabic	சொல் word
1	CV̏,CVC,CV,CV,CV̏ C, CV,CV,CV̏,CV	வீரப்பெருமாள் குருசாமி

3.3. தமிழ் உருபனியல் அமைப்பு

தமிழ் ஒரு திராவிட மொழி. இம்மொழி வினைச்சொல்லை இறுதியாகக் கொண்ட வாக்கியங்களை உடையதாகும். இம்மொழி இலக்கணச் செயற்பாங்குகளை (அதாவது இலக்கணப் பொருண்மைகளை) உருபனியல் நிலையில் வெளிப்படுத்தும்

ஒட்டுக்கள் நிறைந்த மொழியாக விளங்குகிறது. இம்மொழி உள்ள ஒவ்வொரு வேர் உருபனும் ஆயிரக்கணக்கான சொற்களின் உருவாக்கத்திற்கு (அதாவது திரிபு மற்றும் ஆக்க வடிவுகளின் உருவாக்கத்திற்கு) அடிப்படையாக அமைகிறது. தமிழ் வாக்கியங்கள் எழுவாய்தொடரையும் வினையையும் கொண்ட அமைப்பாகவோ எழுவாய்த் தொடரையும் செயப்படுபொருள் தொடரையும் வினையும் கொண்ட அமைப்பாகவோ அமையும்.

3.3.1 பெயர்ச் சொற்களின் உருபனியல் அமைப்பு

தமிழ்ச் சொற்களை உருபனியல் அடிப்படையிலும் தொடரியல் அடிப்படையிலும் பெயர்ச்சொல், வினைச்சொல், பெயரடை, வினையடை என்ற முக்கியமான சொல் வகுப்புகளாகப் பகுப்பர். இவற்றில் பெயர்ச்சொற்கள் வேற்றுமை உருபுகளும் பன்மையை உணர்த்தும் ஒட்டும் ஏற்று வரும். மேலும் பெயர்கள் *ஆக*, *ஆய்*, *ஆன* என்ற ஒட்டுக்களையும் ஏற்று வரும்.

3.3.1.1 பெயரின் திரிபு

பெயர்கள் வேற்றுமைக்கும் எண்ணிற்கும் திரிபுறும் எனப் பார்த்தோம். திரிபுற்ற ஒரு பெயர்ச்சொல் பெயர்ப்பகுதி, எண் மற்றும் வேற்றுமை என்ற

மூன்று உருபன்களின் ஒலியனியல் உட்படுத்தம் ஆகும். இதைப் பிவருமாறு வெளிப்படுத்தலாம்:

பெயர் (+ எண்) (+ வேற்றுமை)

ஒலியனியல் நிலையில் திரிபுற்ற பெயர் வடிவம் மேற்சொன்ன மூன்று உருபன்களையும் மெய்ப்படுத்தும் செய்யும் மூன்று உருபுகளை மட்டும் எப்பொழுதும் கொண்டிருப்பதில்லை. ஏனென்றால் பெயர்த் திரிபுன் போது இலக்கண உருபன்களை மெய்ப்படுத்தம் செய்யும் எண் மற்றும் வேற்றுமைப் பின்னொட்டுகள் தவிர சொல் அல்லது இலக்கணச் செயல்பாடில்லாத சில ஒலியனியல் ஒட்டுக்களும் வரலாம். இவ்வொலியனியல் பின்னொட்டுகள் வெற்றுருபன்கள் (Empty morpheme) எனப்படும். அதாவது அவை ஒரு உருபனை உருப்படுத்தம் செய்யாது. இலக்கண நூல்களில் இவ்வுருபனியல் பின்னொட்டுகள் சாரியைகள் (inflectional increments) எனப்படும். கட்டாய மற்றும் விருப்பு ஒலியனியல் பின்னொட்டுகள்/சாரியைகள் இவற்றுக்கிடையில் வேறுபாட்டைக் கருதலாம். நாம் கட்டாயத் திரிபு சாரியைக்ளைத்

திரிபு (oblique) சாரியைகள் அல்லது திரிபு ஒட்டுகள் என்றும் கூறலாம். விருப்பு திரிபு சாரியைகளை ஒலி நிரவல் (euphonic) சாரியை அல்லது ஒலி நிரவல் பின்னொட்டுகள் என்றும் குறிப்பிடலாம்.

உள்ளுறை அமைப்பில் இரு இலக்கண உருபன்கள் (ஒன்று எண் ஆகவும் மற்றொன்று வேற்றுமையாகவும் வரும்) பெயர் உருபனுடன் வரும். ஒலியனியல் நிலையில் நான்கு வகையான உருபுகள் (பின்னொட்டுகள்) பெயர்ப் பகுதியுடன் சேர்ந்து வரும். அவை பின்வருவனவாகும்.

1. பன்மைப் பின்னொட்டு
2. திரிபு பின்னொட்டு (சாரியை)
3. ஒலிநிரவல் பின்னொட்டு (சாரியை)
4. வேற்றுமைப் பின்னொட்டு

இந்நான்கு பின்னொட்டுகளும் பெயர்ப்பகுதியுடன் பின்வருமாறு வரும்:

1. பெயர் எண்ணுக்காகத் திரிபுறும்போது பன்மைப் பின்னொட்டு பெயர்ப் பகுதியில் சேர்க்கப்படும்.

   பெயர்ப் பகுதி + பன்மைப் பின்னொட்டு

   எ.கா.

   வீடு- கள், எலி- கள்

2. பெயர் வேற்றுமைக்காகத் திரிபுறும்போது வேற்றுமைப் பின்னொட்டு பெயர்ப் பகுதியுடன் நேரிடையகவோ திரிபுப் பகுதியுடனோ சேர்க்கப்படும்.

பெயரின் திரிபுப் பகுதி பெயருடன் திரிபு ஒட்டைச் சேர்ப்பதனாலோ நிலை மொழியின் இறுதி மெய்யொலியை இரட்டுவதாலோ உருவாக்கப்படும். ஒலி நிரவல் சாரியை *இன்* (சில நேர்வுகளில் –*அன்*) பெயர் அல்லது திரிபுப் பகுதிக்கும் வேற்றுமைப் பின்னொட்டுக்கும் இடையில் விருப்பாக வரும். இவ்வுருபன் அமைப்பொழுங்கைப் (morphotactics) பின்வருமாறு தரலாம்:

பெயர்ப்பகுதி / திரிபுப்பகுதி + (ஒலிநிரவல் சாரியை) + வேற்றுமைப் பின்னொட்டு

வீடு- கள்(-இன்)- ஐ > வீடுகளை/வீடுகளினை

கை-கள்(-இன்)- ஐ > கைகளை/கைகளினை

பெயர் அல்லது திரிபுப் பகுதியுடன் பன்மை, வேற்றுமை மற்றும் ஒலியன் ஒட்டுகளைச் சேர்க்கும்போது பல உருபொலியனியல் விதிகள் செயல்படும். எடுத்துக்காட்டாக இடைவெளியைத் (hiatus) தவிர்க்க *வ்* அல்லது *ய்* என்ற ஒலியன் வடிவு உடன்படுத்தியாகச் (glide) சொருகப்படும். இதைப்பற்றி பின்னால் விரிவாகக் கூறப்பட்டுள்ளது.

3.3.1.2 பெயர்ப் பகுதி

பெயர்ப்பகுதி என்பது திரிபுப் பின்னொட்டுகள் சேர்க்கப்படாத பெயர் வடிவமாகும். இவ்வடிவில்தான் பெயர்கள் அகராதியில் பட்டியலிடப் பட்டுள்ளன. இவை எழுவாய் வடிவிலும் வரும். பெயர்ப்பகுதி தனியாகவோ கலப்பாகவோ வரலாம். பெயர்ப்பகுதி ஒரு பெயர் வேரைக் கொண்டிருக்கும் (எ.கா. *கல்*); ஒரு கலவைப் பெயர்ப்பகுதி ஒரு வேரையும் (எ.கா. ஒரு வினைப் பகுதியும்) ஒரு ஆக்க ஒட்டையும் (எ.கா. பெயராக்கப் பின்னொட்டு *ப்பு*) கொண்டிருக்கலாம்.

[பெ[வி படி] -ப்பு]

அல்லது அது ஒரு வேரையும் பால் பின்னொட்டையும் கொண்டிருக்கலாம்.

மாண-வ்-அன்>மாணவன்

மாண-வ்+இ > மாணவி

மாண-வ்-அர்> மாணவர்

தலை-வ்-அன்> தலைவன்

தலை-வ்-இ> தலைவி

தலை-வ்-அர் > தலைவர்

இதில் -*அன்* என்பது ஆண்பால் பின்னொட்டாகும். -*இ* என்பது பெண்பால் பின்னொட்டாகும். -*அர்* என்பது பலர்பால் (உயர்வுப் பன்மை) பின்னொட்டாகும். பல பெயர்ப்பகுதிகள் *அ* என்பதில் முடியும் வேரையும் பகுதியை உருவாக்கும் (stem forming) பின்னொட்டான *ம்* என்பதையும் கொண்டிருக்கும்.

மர-ம் > மரம்

பழ-ம் > பழம்

நில-ம்> நிலம்

திட்ட-ம்> திட்டம்

3.3.1.3 திரிபுப் பகுதி

சில பெயர்கள் பெயர்ப்பகுதியுடன் திரிபுப்பகுதி (oblique stem) என்று அழைக்கப்படும் திரிபுப் பெயர் வடிவத்தையும் கொண்டிருக்கும். சில திரிபுப் பகுதியைக் கொண்டிருக்காது. திரிபுப்பகுதிக்கு நான்கு மடங்கு செயல்பாடும் சூழல் வருகையும் இருக்கும். மேற்சொன்ன எடுத்துக்காட்டில் காட்டியுள்ளது போல் முதலாவது, இப்பெயர் வடிவத்துடன் தான் வேற்றுமைப்பினொட்டுகள் சேர்க்கப்படுகின்றன; இரண்டாவது பெயரின் வடிவம்தான் பெயர் ஒரு அடையாகக் கிழமைச் செயல்பாட்டுடன் தலைமைப் பெயருக்கு முன்னால் வரும் வடிவமாகும்; மூன்றாவது இப்பெயர் வடிவம்தான் சூட்டுப்பெயர் கட்டுமானத்தில் முதல் பெயர்த் தனிமமாக வரும் வடிவமாகும். நான்காவது இப்பெயர் வடிவம்தான் பல பின்னுருபுகளுக்கு முன்வரும் வடிவமாகும். திரிபுப்பகுதியாக வரும். பெயர்கள் திரிபுப் பகுதியாக வருமா வராதா என்பதன் அடிப்படையிலும் திரிபு அடியாக வந்தால் அவற்றின் ஒலியனியல் வடிவம் அடிப்படியிலும் பெயர்களை நான்கு வகுப்புகளாக வேறுபடுத்திக் காட்டலாம்.

3.3.1.3.1. த்து என்ற திரிபு ஒட்டு உடைய திரிபுப் பகுதி

இவ்வகுப்பில் வரும் பெயர்கள் *அம்* என்ற இறுதி அசையைக் கொண்டிருக்கும் (எ.கா. மரம், நிலம், பழம்). இப்பெயர்கள் அவற்றின் இறுதி ஒலியனான *ம்* என்பதை *த்து* என்ற திரிபு ஒட்டால் இடம்பெயர்த்துத் திரிபுப் பகுதியை உருவாக்கும்.

மர-த்து > மரத்து, பழ-த்து>பழத்து, நில-த்து>நிலத்து, தோட்ட- த்து>தோட்டத்து

இத்திரிபுப்பகுதியின் வருகையின் நான்கு சூழல்கள் கீழே எடுத்துக் காட்டப்பட்டுள்ளன.

1. திரிபுப் பகுதி + வேற்றுமை ஒட்டு

   தோட்ட-த்த்-இல் > தோட்டத்தில்

2. திரிபுப்பகுதி + தலைமைப் பெயர்

   தோட்ட-த்து-க் கதவு > தோட்டத்துக் கதவு

3. திரிபுப் பகுதி + பெயர் (பெயர்-பெயர் கூட்டில்)

   தோட்ட- த்து- ப் பூ > தோட்டத்துப் பூ

4. திரிபுப் பகுதி- பின்னுருபு

தோட்ட-த்து-ப் பக்கம் > தோட்டத்துப் பக்கம்

திரிபு வடிவுகளைத் தொடர்ந்து வரும் பெயர்களின் அல்லது பின்னுருபுகளின் தொடக்க அடைப்பு ஒலிகளாக *க், ச், த், ப்* என்பனவற்றின் அடிப்படையில் முன்னர் வரும் திரிபுப் பகுதி சம்பந்தப்பட்ட அடைப்பொலிக்கு இரட்டித்த ஒலியைக் கொண்டிருப்பதை உணராலாம்.

3.3.1.3. 2. அற்று என்ற திரிபு ஒட்டுடைய திரிபுப் பகுதி

இவ்வகுப்பு ஐந்து பெயர்களைக் (மாற்றுப் பெயர்களை) கொண்டிருக்கும். இவை ஒரே திரிபு ஒட்டை ஏற்கும் என்பதன் அடிப்படையில் பொதுவான பண்பைக் கொண்டிருக்கின்றன. இவை பல, சில,அவை மற்றும் எல்லாம் என்பனவாகும். இப்பெயர்கள் அவற்றின் திரிபுப் பகுதியை *அற்று* என்ற திரிபு ஒட்டு சேர்க்கப்படும்போது உருவாக்கும். எல்லாம் என்பதில் *ம்* என்பது நீக்கப்படும். *இவை*, *அவை* என்பனவற்றில் *ஐ* என்பது *அ* என சுருக்கப்படும்.

பல-வ்-அற்று > பலவற்று

அ-வ்-அற்று > அவற்று

எல்லா-வ்-அற்று > எல்லாவற்று

இந்த வகையில் படும் திரிபுப் பகுதி வரும் சூழல்கள் கீழே பட்டியலிடப்பட்டுள்ளன.

1. திரிபுப் பகுதி + வேற்றுமை ஒட்டு

   பல-வ்-அற்-ஓடு > பலவற்றோடு

2. திரிபுப் பகுதி + தலைமைப் பெயர்

   பல- வ்- அற்-இன் விலை > பலவற்றின் விலை

3. திரிபுப் பகுதி + முன்னுருபுத் தொடர்

   பல-வ்- ற்-இன் பக்கம் > பலவற்றின் பக்கம்

இவ்வகையிற்படும் பெயர்களுடன் (மாற்றுப்பெயர்களுடன்) *இன்* என்ற ஒலிநிரவல் சாரியை கட்டாயமாக அற்று என்ற திரிபு ஒட்டுக்குப் பின்னால் அவை பெயர் விசேடணமாக வரும்போதொ அல்லது பின்னுருபால் தொடரப்படும்போதோ வரும்.

3.3.1.3.3 மெய்யொலிகள் இரட்டிப்பால் உருவாக்கப்படும் திரிபுப் பகுதி

டு அல்லது *று* என்ற அசைகளால் இறுதியுறும் எல்லாப் பெயர்களும் இவ்வகுப்பில் அடங்கும் (எ.கா. *வீடு*, *நாடு*, *ஆறு*, *கிணறு*). இவ்வகுப்பைச் சார்ந்த பெயர்கள் இறுதி அசையின் மெய்யெழுத்துக்கள் இரட்ட திரிபு அடியை உருவாக்கும்.

வீடு > வீட்டு

நாடு > நாட்டு

ஆறு > ஆற்று

கிணறு > கிணற்று

இவ்வகையிற்படும் திரிபு அடியின் வருகைச் சூழல் கீழே பட்டியலிடப்பட்டுள்ளன.

1. திரிபுப் பகுதி + வேற்றுமை ஒட்டு

   வீட்- ட்-இல்> வீட்டில்

   கிணற்- ற்- இல்>கிணற்றில்

2. திரிபுப் பகுதி + தலைமைப் பெயர்

   வீட்-டு-க் கதவு>வீட்டுக் கதவு

   கிணற்- று-ச் சுவர்>

3. திரிபுப் பகுதி + பெயர்

   வீட்-டு வேலை > வீட்டுவேலை

   கிணற்-று-த் தண்ணீர் > கிணற்றுத் தண்ணிர்

4. திரிபுப் பகுதி + முன்னுருபுத் தொடர்

   வீட்-டு-ப் பக்கம் >வீட்டுப் பக்கம்

   கிணற்-று-ப் பக்கம் >கிணற்றுப் பக்கம்

மேற்சொன்ன மூன்று வகுப்பில் படாத பெயர்ச் சொற்கள் யாவும் பெயரின் நான்காவது வகுப்படிச் சாரும். இவைகள்தாம் பெயர்களில் பெரும்பான்மையாகும் பிற வகுப்பைச் சார்ந்த பெயர்கள் அவற்றின் திரிபுப் பகுதியாக வருமிடங்களில் இவை அவற்றின் பெயர்ப் பகுதியாக வரும்.

1. திரிபுப் பகுதி + வேற்றுமை ஒட்டு

   தலை-ய்-இல் > தலையில்

2. திரிபுப் பகுதி + தலைமைப் பெயர்
   தலை மயிர்
3. திரிபுப் பகுதி + பெயர்
   தலை வலி
4. திரிபுப் பகுதி + பின்னுருபு
   தலை மேல்

3.3.1.3 4 மாற்றுப் பெயர்களின் திரிபுப் பகுதி

மாற்றுப் பெயர்கள் பெயர்ச்சொல்லின் வகைப்பாட்டின் ஒரு அடைப்பட்ட துணைக் குழுமமாகும். மாற்றுப் பெயர்கள் பெயர்களின் துணைக்குழுமமாகக் குழுமப்படுவதின் காரணம் பெரும்பாலும் தொடரியல் சார்ந்ததாகும். எனவே மாற்றுப் பெயர்கள் தொடரியல் வகைப்பாட்டைச் சார்ந்ததாகக் கருதப்படிகின்றன. ஐந்து வகைப்படும் மாற்றுப் பெயர்களின் ஒரு வகைக்கு அதாவது மூவிட மாற்றுப் பெயர்களுக்குத் திரிபுப்பகுதி இருக்கின்றது. முதலாவது, இரண்டாவது மற்றும் நான்காவது இடமாற்றுப் பெயர்களும், மூன்றாவது இடப்பன்மை அஃறிணை மாற்றுப் பெயர்களுக்கும் திரிபுப்பகுதி இருக்கின்றது. முதலாவது, இரண்டாவது மற்றும் நான்காவது மாற்றுப் பெயர்கள் அவற்றின் திரிபுப் பகுதியை அவற்றின் மாற்றுப்பெயர்கள் வடிவின் ஒலியனியல் மாற்றத்தால் உருவாக்கின்றது. படர்க்கை இடப் பன்மை அஃறிணை மாற்றுப்பெயர்களுக்கு அற்று திரிபு ஒட்டுபெற்ற திரிபுப்பகுதி இருக்கின்றது. பின்வரும் அட்டவணை மாற்றுப் பெயர்களையும் அவற்றின் திரிபு வடிவங்களையும் காட்டும்.

இடம்	**மாற்றுப்பெயர்**	**திரிபு வடிவம்**
தன்மை ஒருமை	நான்	என் (ன்), என
முன்னிலை ஒருமை	நீ	உன் (ன்), உன்
படர்க்கை ஒருமை ஆண்பால்	இவன்/அவன்	இவன்/அவன்
படர்க்கை ஒருமை பெண்பால்	இவள்/அவள்	இவள்/அவள்
படர்க்கை ஒருமை அஃறிணை	இது/அது	இது/அது

படர்க்கை ஒருமை ஆண்பால் / பெண்பால்	இவர்/அவர்	இவர்/அவர்
தற்சுட்டு ஒருமை	தான்	தன்/தன
தன்மைப் பன்மை	நாங்கள்/நாம்	எங்கள்/ நம்(ம்), நம
முன்னிலைப் பன்மை	நீங்கள்	உங்கள்
படர்க்கைப் பன்மை ஆண்பால்/ பெண்பால்	இவர்கள்/அவர்கள்	இவர்கள்/அவர்கள்
படர்க்கை பன்மை/ அஃறிணை	இவை/அவை இவைகள்/அவைகள்	இவற்று/அவற்று இவைகள்/அவைகள்
தற்சுட்டுப் பன்மை	தாங்கள்	தங்கள்

மேற்சொன்ன அட்டவணைப்படி *நான்* என்ற தன்மைப் பெயருக்கு மூன்று திரிபுப் பகுதி வடிவங்கள் இருக்கின்றன: *என், என்ன், என.* *என* நான்காவது வேற்றுமைப் பின்னொட்டுக்கு முன்னும், *என்ன்* என்பது *இடம்* என்ற பின்னுருபுக்கு முன்னும், *என்* தலைமைப் பெயருக்கு முன்னும் பிற பின்னுருபுகளுக்கு முன்னும் வரும்.

1. திரிபுப் பகுதி + வேற்றுமை ஒட்டு
   என-க்கு > எனக்கு, என்ன்-இடம் > என்னிடம்
2. திரிபுப் பகுதி + தலைமைப் பெயர்
   என் வீடு
3. திரிபுப் பகுதி + பின்னுருபு
   என் பக்கம்

3.3.1.3 .5 எண் குறிப்பு

தமிழில் ஒருமை, பன்மை என்ற இரண்டு எண்கள் இருக்கின்றன. பன்மை ஒட்டு *-கள்/-க்கள்* என்பதால் உருப-ஒலியனியல் அடிப்படையில் குறிப்பீடு செய்யப்பட்டுள்ளது. முன்னர் கூறியபடி பன்மை ஒட்டு பெயர்ப்பகுதியுடன் சேர்க்கப்படுகிறது; பெயரின் திரிபு அடியுடன் சேர்க்கப்படுவதில்லை. பன்மை ஒட்டிற்கு *கள், க்கள்* என்ற இருமாற்றுருபுகள் இருக்கின்றன. அவற்றின் வருகை பின்வருமாறு அமையும்: *க்கள்* நெட்டுயிர் இறுதியில்

வரும் எல்லாப் பெயர்களுடனும் இரண்டு குற்றுயிர் உள்ள அசைகளைக் கொண்ட உ என்ற குற்றுயிரால் இறுதியுறும் பெயர்களுடனும் சேர்ந்து வரும்; பிற இடங்களில் *கள்* வரும்.

ஒருமைப் பெயர் -		பன்மைப் பெயர்	ஒருமைப்பெயர் -		பன்மைப் பெயர்
பூ	-	பூ- க்கள்	கண்	-	கண்- கள்
ஈ	-	ஈ- க்கள்	மரம்	-	மரங்- கள்
பசு	-	பசு- க்கள்	வீடு	-	வீடு-கள்
தெரு	-	தெரு- க்கள்	கல்	-	கற்- கள்

3.3.1.3.6 ஒலி நிரவல் சாரியைகள் *இன், அன்*

முன்னர் *த்து*, *அற்று* என்ற திரிபுச் சாரியைகளைப் பற்றிப் பார்த்தோம். இவை தவிர *இன்*, *அன்* என்ற ஒலியனியல் பின்னொட்டுக்கள் வடிவத்திலிருக்கும் விருப்புத் திரிபுச் சாரியைகளும் இருக்கின்றன. இவை ஒலி நிரவல் முன்னொட்டுக்கள் என்று அழைக்கப்பெறும். *இன்* என்பது எல்லாப் பெயர்களுடன் மட்டும் வரும். *அன்* என்பது *இது* மற்றும் *அது* என்ற மாற்றுப்பெயர்களுடன் மட்டும் வரும். இவ்வொலி நிரவல் ஒட்டுக்களான *இன்* என்பதும் *அன்* என்பதும் வரும் சூழல்கள் கீழே தரப்பட்டுள்ளன:

1. வேற்றுமைக்காகத் திரிபுற்ற பெயர்களில் ஒலிநிரவல் பின்னொட்டு பெயர்ப்பகுதிக்கும் அல்லது திரிபுப் பகுதிக்கும் அல்லது பன்மை ஒட்டுக்கும் வேற்றுமைப் பின்னொட்டுக்கும் இடையில் வரும்.

கை(-இன்)-ஆல் > கையால்/கையினால்

அத்(-அன்)-ஆல் > அதால்/அதனால்

யுத்த-த்த்(-இன்)-ஆல் > யுத்தத்தால்/யுத்தத்தினால்

கற்-கள்(-இன்)-ஆல் >கற்களால்/கற்களினால்

2. அடையாக வரும் பெயர்களுடன் உடைமை வேற்றுமைச் செயல்பாட்டில் தன்மைப் பெயருக்கு முன் வரும்.

கை(-ய்-இன்) புத்தகம் > கையின் புத்தகம்

மர(-த்த்-இன்) கிளை > மரத்தின் கிளை

பையன்-கள்(-இன்) சட்டை > பையன்களின் சட்டை

3. ஒரு கிளமைப் பின்னொட்டுக்கு முன் விருப்பாக, சில பெயர்களுடன் கட்டாயமாகப் பெயர்ப்பகுதிக்கு அல்லது திரிபுப்பகுதிக்கு அல்லது பன்மை ஒட்டுக்குப் பின் வரும்.

சாவி(ய்-இன்) மூலம் > சாவியின் மூலம்

அத்(-அன்) மூலம் > அதன் மூலம்

மர(-த்த்(-இன்) மூலம் > மரத்தின் மூலம்

ஆணி-கள்(-இன்) மூலம் >ஆணிகளின் மூலம்

3.3.1.3.7 வேற்றுமைக் குறியீடு

பெயர் அல்லது பெயர்த்தொடருக்கும் வினைக்கும் அல்லது இரண்டு பெயருகளுக்கும் அல்லது பெயர்த்தொடர்களுக்கும் இடையில் வரும் தொடரியல் மற்றும் பொருண்மையியல் உறவுகளை வேற்றுமை வெளிப்படுத்தும். வினைப் பயனிலைகள் ஒரு வாக்கியத்தில் கட்டாய மற்றும் விருப்புப் பங்கெடுப்பாளர்களை எடுக்கும் (எ.கா. எழுவாய்ப் பெயர்த் தொடர்கள், செயப்படுபொருள், பெயர்த்தொடர்கள், வினையடைகள்). இவ்வாறு வேற்றுமைக் குறியீடுகள் பெயர்த்தொடருக்கும் வினைப் பயனிலைக்கும் இடையிலுள்ள தொடரியல் மற்றும் பொருண்மையியல் உறவை வெளிப்படுத்துகிறது. அதாவது அவை ஒரு பங்கெடுப்பாளர் பெயர்த் தொடரின் இலக்கணச் செயன்மையைக் காட்டும் (எ.கா. எழுவாய் செயல்படு பொருள்). மேலும் அவை பயனிலையாக்கத்தில் பங்கெடுப்பாளர் பெயர்த்தொடரின் பொருண்மையியல் பங்களிப்பை வெளிப்படுத்தும் (எ.கா. ஒரு செயலின் அல்லது நிகழ்வின் காரணம், நிகழ்வு அல்லது நிலையின் இடம், இயக்கத்தின் மூலம் அல்லது இலக்கு). உடைமை வேற்றுமை இரண்டு பெயர்த்தொடருக்கு இடையில் உள்ள உறவை வெளிப்படுத்தும்.

உருபனியல், தொடரியல் அடிப்படையில் வேற்றுமைக் குறியீடு இரு வழிகளில் உருப்படுத்தம் செய்யப்படுகிறது.

1. வேற்றுமைப் பின்னொட்டுகளால்
2. பின்னுருபுகளால் (இவற்றில் சில கட்டுண்ட வடிவுகள்)

பின்னொட்டுக்களைப் பொறுத்தவைரையில் வேற்றுமைக் குறிப்பு வேற்றுமை பின்னொட்டு மற்றும் பின்னுருபு இவற்றின் ஒன்றுசேர்க்கையால் உருப்படுத்தம் செய்யப்படுகிறது. எடுத்துக்காட்டாகக் கருவியின் பொருணமைப் பங்களிப்பு *ஆல்* என்ற பின்னொட்டாலும் (எ.கா. *சாவியால்*) மூலம் என்ற பின்னுருபாலும் (எ.கா. *சாவி மூலம்*) வெளிப்படுத்தலாம் இறுதியாகக் கருவியின் பொருண்மையியல் பங்களிப்பை ஒரு வேற்றுமைப் பின்னொட்டை ஆளும் பின்னுருபால் வெளிப்படுத்தலாம். கொண்டு என்ற பின்னுருபு இரண்டாம் வேற்றுமை உருவான *ஐ*-யை ஆளுகை செய்யும் (எ.கா. *சாவியைக் கொண்டு*). தற்காலத் தமிழில் ஆறு வேற்றுமைப் பின்னொட்டுகள் இருக்கின்றன.

வேற்றுமை	**பின்னொட்டுகள்**
செயப்படுபொருள்	ஐ
கொடைவேற்றுமை	(உ)க்கு
கருவி வேற்றுமை	ஆல்
உடன்நிகழ்ச்சி வேற்றுமை	ஓடு
இட வேற்றுமை	இல்
வேற்றுமை	அது

ஐந்து கட்டுண்ட பின்னுருபுகள் இருக்கின்றன.

கட்டுண்ட வேற்றுமை	**கட்டுண்ட பின்னுருபுகள்**
இடவேற்றுமை	இடம்
உடன்நிகழ்ச்சி வேற்றுமை	உடன்
மூல வேற்றுமை	இருந்து
பயனடை வேற்றுமை	ஆக

மரபு இலக்கணவியலார் வேற்றுமைப் பின்னொட்டுகளையும் *ஆக* தவிர பிற கட்டுண்ட பின்னுருபுகளையும் பல இடப் பின்னுருபுகளையும் வேற்றுமை குறிகளாக எடுத்தாண்டனர். இக்குறிகளின் அடிப்படையில் அவர்கள் எட்டு வேற்றுமைகளை வேறுபடுத்தினர். இவைகள் அவற்றின் ஒலியனியல் வடிவத்தாலோ வரிசை எண்ணாலோ அடையாளப்படுத்தப்பட்டன. வேற்றுமைக் குறியீட்டால் குறிப்பிடப்படாத விளி

வேற்றுமையும் அவர்கள் சேர்த்துக்கொண்டனர். பின்னர் மேனாட்டு இலக்கணத்தார்கள் இலத்தின் அல்லது சமஸ்கிருத இலக்கணக் கலைச்சொற்களைப் பயன்படுத்தி வேற்றுமை குறியின் செயல்பாடு அடிப்படையில் எட்டு வேற்றுமைகளை புலக்குறிப்பு செய்தனர்

1. எழுவாய் வேற்றுமை
2. செயல்பாடு பொருள் வேற்றுமை
3. கொடை வேற்றுமை
4. கருவி வேற்றுமை
5. உடன்நிகழ்ச்சி வேற்றுமை
6. இட வேற்றுமை
7. நீங்கல் வேற்றுமை
8. உடைமை வேற்றுமை
9. விளி வேற்றுமை

வேற்றுமைக் குறிகளின் மேற்சொன்ன புலக்குறிப்புகள் ஒரு வேற்றுமைக் குறியின் பொருண்மையை முழுவதுமாக உருப்படுத்தம் செய்யவில்லை; ஏனென்றால் அவை அவற்றின் செயல்பாடுகளில் ஒன்றைத்தான் வெளிப்படுத்துகின்றது. எடுத்துக்காட்டாகக் கருவி வேற்றுமைப் பின்னொட்டு என்று புலக்குறிப்பு செய்யப்பட்ட *ஆல்* என்ற வேற்றுமைப் பின்னொட்டு கருவிச் செயல்பாட்டை வெளிப்படுத்துவதுடன் பிற கருத்துக்களையும் வெளிப்படுத்துகின்றன.

வேற்றுமை அடையாளக் குறியீட்டைப் பொருத்தவரையில் பின்வரும் பாகுபாட்டு அடிப்படையில் விளக்கலாம்:

1. வேற்றுமைப் பின்னொட்டு
2. கட்டுண்ட பின்னுருபுகள்
3. பின்னுருபுகள் (சுதந்திரமான வடிவுகள்)

பின்வரும் அட்டவணை வேற்றுமை உருபுகளையும் அவற்றின் வருகை முறைகளையும் அவற்றின் வெறுபட்ட செயல்பாடுகளையும் தரும்:

வேற்றுமை	**பின்னொட்டு/**	**வருகை முறை**	**பொருள்/ செயல்பாடு**

	பின்னுருபு		
1. எழுவாய் வேற்றுமை	-Ø		**1. எழுவாய்** எ.கா. *ராமன் வந்தான்* **2. பயனிலை** எ.கா. *குமரன் மாணவன்* **3. எழுவாய் நிரப்பி** எ.கா. குமரன் *மாணவன்* ஆவான் 4. **செயப்படுபொருள் நிரப்பி** எ.கா. குமரன் *ராமனைத் தலைவன்* ஆக்கினான் 5. **செயப்படுபொருள்** எ.கா. குமார் *தண்ணீர்* கேட்டான்
2. செயப்படு பொருள் வேற்றுமை	-ஐ		**1. செயப்படு பொருள்** எ.கா. ராமன் பையனை பார்த்தான்
3. கொடை வேற்றுமை	-க்கு~ -உக்கு~-கு	1. பெயர்களின் மற்றும் மாற்றுப் பெயர்களின் திரிபுப் பகுதியுடன் -க்கு வரும். எ.கா. மரத்து-க்கு பலவற்று-க்கு, ஆற்று-க்கு, என-க்கு 2. இ/ஈ/ஐ/அய் என்று முடியும் பெயரடிகளுடன்	1. **அயற்செயப்படு பொருள்** எ.கா. குமார் *அப்பாவுக்கு* ஒரு படத்தைக் காட்டினான் 2. **இயக்கத்தின் இலக்கு** எ.கா. குமார் *அப்பாவுக்கு* ஒரு கடிதம் கொடுத்தான் 3. **கார்ணம்** எ.கா. குமார் *பணத்துக்குத்தான்* வேலை செய்கிறான் 4. **அனுபவிப்பவர்** எ.கா. *குமாருக்கு* ஒரு புத்தகம்

		க்கு வரும் எ.கா. தம்பிக்கு, ஈக்கு, சட்டைக்கு நாய்க்கு, படிப்புக்கு 3. பிற எல்லாப் பெயரடிகளுக்குப் பின் *உக்கு* வரும். எ.கா. அம்மா-உக்கு, பால்- உக்கு, தெரு-வ்-உக்கு, பையன்-உக்கு, பெண்ண்-உக்கு 4. கு ஒலிநிரவல் சாரியை *இன்*, *அன்* என்பனவற்றைத் தொடர்ந்து வரும் எ.கா. பசு+இன்+கு> பசுவிற்கு, அது+அன்+கு அதற்கு	இருக்கிறது. 5. **காலத்தின் கூறு** எ.கா. குமார் *ஐந்து மணிக்கு* வருவான் 6. **விகிதம்** எ.கா. மருந்து ஒரு *நாளைக்கு* மூன்று வேளை சாப்பிடு 7. **பகிர்ந்தளிப்புச் செயல்பாடு** எ.கா. *ஆளுக்கு* ஒரு டீ போடு 8. **ஒப்புமை மாதிரி** எ.கா. *குமாருக்கு* இவன் நல்லவன் 9. **குறிப்பிடும் இடம்** எ.கா. *மதராசுக்கு* நூறு மைல் பின்னால்
4. பயனடை வேற்றுமை	-க்கு-ஆக ~ -உக்கு-ஆக ~ -கு-ஆக	1. க்கு-ஆக எ.கா. நாய்-க்கு-ஆக 2. உக்கு-ஆக எ.கா.அவன்-உக்கு-ஆக 3. கு- ஆக எ.கா.பசுவிற்-கு-ஆக	1. **பயனடைவு** எ.கா. ராஜா தன் பையனுக்காகப் பணம் சேர்த்தான் 2. **காரணம்** எ.கா. ராஜா தன் *அப்பாவுக்காகச்* சீக்கிரம் வீட்டுக்கு வாந்தான்

5. கருவி வேற்றுமை	-ஆல்		1. **கருவி** எ.கா. குமார் கத்தியால் பழத்தை வெட்டினான் 2. **வழி** எ.கா. குமார் தன் கண்ட *அனுபவத்தால்* இதைக் துளைக்கக் கற்றுக்கொண்டான். 3. **மூலம்/ உருபொருள்** எ.கா. குமார் *மண்ணால்* இந்தப் பந்தைச் செய்தான். 4. **காரணம்-1** எ.கா. *வெயிலால்* குமார் வரவில்லை. 5. **காரணம் -2** எ.கா. *மழையால்* பயிர் நன்றாக வளர்ந்தது
6. உடனிகழ்ச்சி வேற்றுமை	-ஓடு/ -உடன்		1. **உடனிகழ்ச்சி** எ.கா. குமார் தன் *மனைவி-ய் -ஓடு/-உடன்* வந்தான் 2. **முறை** எ.கா. குமார் *அன்ப்-ஓடு/*-உடன் சிரித்தான். 3. **வழி** எ.கா குமார் அப்பாவின் *உதவி-ய்-ஓடு/*-உடன் இதைப் படித்தான். 4. **அதிகப்படி** எ.கா. ராஜா நான்கு

			இட்லியோடு இரண்டு தோசையும் சாப்பிட்டான். 5. **இடம்** எ.கா. குமார் *வீட்டோடு* இருக்கிறான். 6. **திசை** எ.கா. குமார் *தெருவோடு* நடந்தான் 7. **உட்படுதல்** எ.கா. சித்திரை *மாசத்தோடு* தமிழ் வருஷம் ஆரம்பிக்கிறது. 8. **வரையறை** எ.கா. இந்த ஊரில் ஒரு *கோயிலோடு* சரி
7. இட வேற்றுமை	-இல்		1. **இடத்தில் இடம்** எ.கா. *குருவி* மரத்தில் உட்காந்திருக்கிறது. 2. **காலத்தின் புள்ளியிடம்** (எதிர்காலத்துடன்) எ.கா. ராஜா *ஒரு வாரத்தில்* வருவான். 3. **கால அளவு** (இறந்த காலத்துடன்) எ.கா. ராஜ ஒருவாரத்தில் இந்த புத்தகத்தைப் படித்தான். 4. **பாங்கு** எ.கா. குமார் தமிழில் பேசினான்

			5. **இடைபடு உறவு** எ.கா. அந்த மாம்பழங்களில் சில் மல்கோவா. அந்த மாணவர்களில் சிலர் நன்றாக படிப்பவர்கள்.
	-இடம்		1. **இயக்கத்தின் இலக்கு** எ.கா. ராஜா *அப்பாவிடம்* வந்தான். 2. **கைமாற்றத்தின் மூலம்** எ.கா. ராஜா *மந்திரியிடம்* ஒரு பரிசு வாங்கினான். 3. **உணர்ச்சியின் இலக்கு** எ.கா. ராஜா *அப்பாவிடம்* கோபப்பட்டான்.
8. நீக்கல் வேற்றுமை	-இல்/-இடம்–இருந்து, இருந்து	1. இட வேற்றுமைக்குத் திரிபுற்ற பெயரைத் தொடர்ந்து வரும் . எ.கா. மரத்தில் – இருந்து பையனிடம் – இருந்து 2. நேரடியாக இடப்பகுதியுடன் வரும். எ.கா. அங்க்- இருந்து மேல்- இருந்து *-இல்* என்ற	**இயக்கத்தின் மூலம்** எ.கா. ராஜா *கோவிலிலிருந்து* வந்தான். **உருபொருளிலிருந்து பிரிதல்** எ.கா. ராஜா தன் அப்பாவை *நெருப்பிலிருந்து*/ ஒரு *புலியிடமிருந்து* காப்பாற்றினான். **தேர்வின் மூலம்** எ.கா. ராஜா *மாணவர்களிலிருந்து* ஒரு நல்லவனைத்

		இடவேற்றுமைப் பின்னொட்டிற்குத் திரிபுற்ற விலங்கினப் பெயர்களுடன் வருகையிலும் -*இடம்* என்ற இட வேற்றுமைப் பின்னொட்டிற்குத் திரிபு விலங்கினப் பெயர்களுடன் வருகையிலும் -*இருந்து* அதன் செயல்பாடுகளில் (பொருள்களில்) வேறுபடும் (பார்க்க அடுத்தப்பக்கத்தில்)	தேர்ந்தெடுத்தான். **தொடர்ச்சியின் தொடக்கம்** எ.கா. *மந்திரியிலிருந்து* எல்லோரும் லஞ்சம் வாங்குகிறார்.
9. உடைமை வேறுமை	-இன், -உடைய, -அது, பெயர்திரிபு அடி	திரிபு வடிவில் பெயர் எ.கா. மரத்துக்கிளை பெயர் + ஒலிநிரவல் – *இன்* எ.கா. *அரசனின்* முடி பெயர் + உடைமை வேற்றுமை *அது* எ.கா.அரசனது முடி பெயர் + இடைநிலை சாரியை + உடைமைப் பின்னொட்டு எ.கா. *நாயினது* வால் பெயரடைப் பின்னொட்டு உடைய எ.கா. *நாயினுடைய* வால்	பெயரடையாகச் செயல்படுகிறது எ.கா. பள்ளிக்கூடம் குமாருடையது.

3.3.1.3.8 வேற்றுமை உருபுகளும் சந்தி விதிகளும்

இராசேந்திரன் (Rajendran, 2000) தமது தமிழ் உருபனியல் ஆய்வி உருவாக்கதிற்கு (Tamil Morphological Analyzer) வேண்டி தமிழ்ப் பெயர்ச்சொற்களைச் சந்திவிதிகளைக் கருத்திற்கொண்டு பாகுபாடு செய்கையில் பின்வரும் சந்திவிதிகளைப் பட்டியலிடுகின்றார். (இராசேந்திரன்,1994:47-56):

1. *ஆ*, *ஓ*, *ஊ* என முடியும் பெயர்களைத் தொடர்ந்து உயிர்த் தொடக்க வேற்றுமைப் பின்னொட்டுகள் (உ.வே.பி.) வரும்போது *வ்* என்ற உடன்படுத்தி இடையில் வரும்.

-ஆ/ஓ/ஊ + உ.வே.பி→-ஆ/ஓ/ஊ- வ்- உ.வே.பி.

எ.கா.

கடா + ஐ > கடாவை

கடா + ஆல் > கடாவால்

கடா + ஓடு >கடாவோடு

கடா + உடன் > கடாவுடன்

கடா + உக்கு >கடாவுக்கு

கடா + இன் > கடாவின்

கடா + இல் > கடாவில்

2. *இ*, *ஈ*, *ஐ* என்று முடியும் பெயர்கள் -*உக்கு* தவிர்த்த பிற வேற்றுமைப் பின்னொட்டுகளை ஏற்கும்போது *ய்* என்ற உடன்படுத்தி ஒலி இடையில் வரும்.

-இ/ஈ/ஐ + உக்கு தவிர்த்த உ.வே.பி. (இ/ஈ/ஐ–ய்-உ.வே.பி)

எ.கா.

எலி + ஐ > எலியை

எலி + ஆல் > எலியால்

எலி + ஓடு > எலியோடு

எலி + உடன் > எலியுடன்

எலி + இன் > எலியின்

எலி + இல் > எலியில்

3. *இ, ஈ, ஐ, ய்* என்று முடியும் பெயர்கள் முவரும் வேற்றுமைப் பின்னொட்டு -*உக்கு* என்பதன் *உ* கெடும்

-இ/ஈ/ஐ/ய் + உக்கு→இ/ஈ/ஐ/ய்-க்கு

எ.கா.

எலி + உக்கு > எலிக்கு

ஈ + உக்கு > ஈக்கு

யானை + உக்கு > யானைக்கு

நெய் + உக்கு > நெய்க்கு

4. *உ*-வில் முடியும் பெயர்களை அவற்றின் சந்திச் செயல்பாடு அடிப்படையில் நான்கு வகைகளாகப் பிரிக்கலாம்

4.1. (C)VV-ட்-உ என்ற அசை அமைப்பக் கொண்ட சொற்களைத் தொடர்ந்து உயிர்த் தொடக்க வேற்றுமைப் பின்னொட்டுக்கள் வரும்போது *உ* கெட, *ட்* இரட்டிக்கும்.

(C)VV ட் – உ + உ.வே.பி →(C)VVட்ட் – உ.வே.பி

எ.கா.

காடு+ஐ > காட்டை

காடு+ஆல் > காட்டால்

காடு+ஓடு > காடோடு

காடு+உடன் > காட்டுடன்

காடு+உக்கு > காட்டுக்கு

காடு+இன் > காட்டின்

காடு+இல் > காட்டில்

4.2. *று*-வில் முடியும் பெயர்களைத் தொடர்ந்து உயிர்த்தொடக்க வேற்றுமைப் பின்னொட்டுக்கள் வரும்போது *உ* கெட *ற்* இரட்டும்.

-று + உ.வே.பி → ற்ற் – உ.வே.பி.

4.3. (C)VCC-உ அல்லது (C)VVC-உ என்ற அசை அமைப்புடைய பெயர்களைத் தொடர்ந்து உயிர்த்தொடக்க வேற்றுமைப் பின்னொட்டுகள் வருகையில் உ கெடும்

(C)VCC-உ + உ.வே.பி. → CVCC-உ.வே.பி

எ.கா.

வண்டு+ஐ > வண்டை

வண்டு+ஆல் > வண்டால்

வண்டு+ஓடு > வண்டோடு

வண்டு+உடன் > வண்டுடன்

வண்டு+உக்கு > வண்டுக்கு

வண்டு+இன் > வண்டின்

வண்டு+இல் >வண்டில்

காது+ ஐ > காதை

காது+ ஆல் > காதால்

காது+ ஓடு > காதோடோடு

காது+ உடன் > காதுடன்

காது+ உக்கு > காதுக்கு

காது+ இன் > காதின்

காது+ இல் > காதில்

4.4 (C)VCV என்ற அசையமைப்பைக் கொண்டு உ-வில் முடியும் பெயர்களைத் தொடர்ந்து உயிர்த் தொடக்க வேற்றுமை ஒட்டுகள் வரும்போது *வ்* என்ற உடன்படுத்தி இடையில் வரும்.

(C)VC-உ + உ.வே.பி.→ (C)VC-உ-வ்-உ.வே.பி.

பசு+ ஐ > பசுவை

பசு+ஆல் > பசுவால்

பசு+உக்கு > பசுவுக்கு

பசு+ ஓடு > பசுவோடு

பசு+உடன் > பசுவுடன்

பசு+இன் > பசுவின்

பசு +இல் > பசுவில்

5. (C)VC என்ற அசை அமைப்புடைய *ண்*, *ல்*, *ள்*, *ன்* என முடியும் பெயர்களைத் (ஓரசைச் பெயர்களை) தொடர்ந்து உயிர் தொடக்க வேற்றுமைப் பின்னொட்டுகள் வர முறையே *ண்*, *ய்*, *ல்*, *ள்*, *ன்* என்பன இரட்டும். (விதிவிலக்கு *ய்* + *உக்கு* > *ய்க்கு*. எ.கா. *நெய்* + *உக்கு* > *நெய்க்கு*)

(C)Vண்/ய்/ல்/ள்/ன் + உ.வே.பி→ (C)Vண்ண்/ய்ய்/ல்ல்/ள்ள்/ன்ன். உ.வே.பி.

கண்+ ஐ > கண்ணை

கண்+ஆல் >கண்ணால்

கண்+உக்கு > கண்ணுக்கு

கண்+ஓடு > கண்ணோடு

கண்+உடன் > கண்ணுடன்

கண்+இன் > கண்ணின்

கண்+உடைய > கண்ணுடைய

கண்+அது > கண்ணது

கண்+இல் > கண்ணில்

நெய்+ ஐ > நெய்யை

நெய்+ஆல் > நெய்யால்

நெய்+உக்கு > நெய்க்கு

நெய்+ஓடு > நெய்யோடு

நெய்+உடன் > நெய்யுடன்

நெய்+இன் > நெய்யின்

நெய் +உடைய > நெய்யுடைய

நெய்+அது > நெய்யது

நெய்+இல் > நெய்யில்

கல்+ ஐ > கல்லை

கல்+ஆல் > கல்லால்

கல்+உக்கு > கல்லுக்கு

கல்+ஓடு > கல்லோடு

கல்+உடன் > கல்லுடன்

கல்+இன் > கல்லின்

கல்+உடைய > கல்லுடைய

கல்+அது > கல்லது

கல்+இல் > கல்லில்

முள்+ ஐ > முள்ளை

முள்+ஆல் >முள்ளால்

முள்+உக்கு >முள்ளுக்கு

முள்+ஓடு >முள்ளோடு

முள்+உடன் > முள்ளுடன்

முள்+இன் > முள்ளின்

முள் + உடைய > முள்ளுடைய

முள்+அது > முள்ளது

முள்+இல் >முள்ளில்

பொன்+ ஐ > பொன்னை

பொன்+ ஆல் > பொன்னால்

பொன்+உக்கு > பொன்னுக்கு

பொன்+ஓடு > பொன்னோடு

பொன்+உடன் > பொன்னுடன்

பொன்+இன் > பொன்னின்

பொன்+உடைய > பொன்னுடைய

பொன்+அது > பொன்னது

பொன்+இல் > பொன்னில்

6. *அம்* ஈற்றுப் பெயர்களைத் தொடர்ந்து உயிர்த்தொடக்க வேற்றுமைப் பின்னொட்டுக்கள் வரும்போது *ம்* கெட *த்த்* என்ற சாரியை இடையில் வரும்.

-அம் + உ.வே.பி.→-அ-த்த்-உ.வே.பி.

மரம்+ஐ > மரத்தை

மரம்+ஆல் > மரத்தால்

மரம்+உக்கு > மரத்துக்கு

மரம்+ஓடு > பொன்னோடு

மரம்+உடன் > மரத்துடன்

மரம்+இன் > மரத்தின்

மரம்+உடைய > மரத்துடைய

மரம்+அது > மரத்தினது

மரம்+இல் > மரத்தில்

1. *அ* ஈற்றுப் பெயர்களான *பல*, *சில* என்பனவற்றைத் தொடர்ந்து உயிர்த்தொடக்க வேற்றுமைப் பின்னொட்டுகள் வரும்போது *அற்ற்* என்ற சாரியை இடையில் வரும்.

பல + உ.வே.பி.(பல-வ்-அற்ற்-உ.வே.பி.

பல+ ஐ > பலவற்றை

பல+ஆல் > பலவற்றால்

பல+உக்கு > பலவற்றுக்கு

பல+ஓடு>பலவற்றோடு

பல+உடன் > பலவற்றுடன்

பல+இன் > பலவற்றின்

பல + உடைய > பலவற்றுடைய

பல + அது > பலவற்றினது

பல + இல் >பலவற்றில்

8. எல்லாம் என்ற பெயரைத் தொடர்ந்து உயிர்த்தொடக்க வேற்றுமைப் பின்னொட்டுக்கள் வரும்போது *ம்* கெட *அற்ற்* என்ற சாரியை இடையில் வரும்.

எல்லாம் + உ.வே.பி.→எல்லா-வ்-அற்ற் – உ.வே.பி.

எல்லாம் + ஐ > எல்லாவற்றை

எல்லாம் +ஆல் > எல்லாவற்றால்

எல்லாம் +உக்கு > எல்லாவற்றுக்கு

எல்லாம் +ஒடு>எல்லாவற்றோடு

எல்லாம் +உடன் > எல்லாவற்றுடன்

எல்லாம் +இன் > எல்லாவற்றின்

எல்லாம் +உடைய > எல்லாவற்றுடைய

எல்லாம் +அது > எல்லாவற்றினது

9. 5.ஆம் 6- ஆம் விதிகளுக்கு உட்படாத *ண்*, *ல்*, *ள்*, *ன்* மற்றும் *ர்*, *ழ்* என முடியும் பெயர்களைத் தொடர்ந்து உயிர்த்தொடக்க வேற்றுமைப் பின்னொட்டுக்கள் வரும்போது அவை எந்தவித மாற்றமுமின்றி இணையும் (விதிவிலக்கு: *ய்* + *உக்கு* > *ய்க்கு*).

(5-ஆம், 6-ஆம் விதிக்கு உட்படாத) *-ண்/ ல்/ ள்/ ன்/ ர்/ ழ்* + உ.வே.பி. (ண்/ ல்/ ள்/ ன்/ர்/ ழ் + உ.வே.பி.

எ.கா

தூண்+ ஐ > தூணை

தூண்+ஆல் >துணால்

தூண்+உக்கு > தூணுக்கு

தூண் +ஓடு >தூணோடு

தூண் + உடன் > தூணுடன்

தூண்+இன் > தூணின்

தூண்+உடைய >தூணுடைய

தூண்+அது > தூணது

தூண்+இல் > தூணில்

நாய்+ ஐ > நாயை

நாய்+ஆல்>நாயால்

நாய்+உக்கு>நாய்க்கு

நாய்+ஓடு>நாயோடு

நாய்+உடன் > நாயுடன்

நாய்+இன் > நாயின்

நாய் +உடைய > நாயுடைய

நாய் +அது > நாயது

நாய் +இல் > நாயில்.

கால் + ஐ > காலை

கால் +ஆல் > காலால்

கால் +உக்கு > காலுக்கு

கால் +ஓடு > காலோடு

கால் +உடன் > காலுடன்

கால் +இன் > காலின்

கால்+உடைய > காலுடைய

கால்+அது > காலது

கால்+இல் > காலில்.

வாள்+ ஐ > வாளை

வாள்+ஆல்>வாளால்

வாள்+உக்கு > வாளுக்கு

வாள்+ஓடு >வாளோடு

வாள்+உடன் > வாளுடன்

வாள்+இன் >வாளின்

வாள்+உடைய > வாளுடைய

வாள்+அது>வாளது

வாள்+இல்>வாளில்.

மான்+ ஐ > மானை

மான்+ஆல்>மானால்

மான்+உக்கு >மானுக்கு

மான்+ஓடு>மானோடு

மான்+உடன் >மானுடன்

மான்+இன்>மானின்

மான்+உடைய >மானுடைய

மான்+அது>மானது

மான்+இல் > மானில்.

சுவர்+ ஐ > சுவரை

சுவர்+ ஆல் > சுவரால்

சுவர்+உக்கு > சுவருக்கு

சுவர்+ஓடு > சுவரோடு

சுவர்+ உடன் > சுவருடன்

சுவர்+ இன் > சுவரின்

சுவர்+ உடைய > சுவருடைய

சுவர்+ அது > சுவரது

சுவர்+இல்>சுவரில்.

இதழ்+ ஐ > இதழை

இதழ்+ஆல்>இதழால்

இதழ்+உக்கு >இதழுக்கு

இதழ்+ஓடு >இதழோடு

இதழ்+உடன் >இதழுடன்

இதழ்+இன் >இதழின்

இதழ்+உடைய >இதழுடைய

இதழ்+அது>இதழது

இதழ்+இல்>இதழில்.

10. *இன்* என்ற சாரியை ஒருமைப்பன்மைப் பெயர்ப்பகுதிகளைத் தொடர்ந்து வேற்றுமை உருபுகளுக்கு முன்னர் விருப்பாக வரும்.

எ.கா.

பசு + ஐ >பசுவினை

பசு + ஆல் >பசுவினால்

பசு + உக்கு >பசுவினுக்கு

பசு + ஓடு >பசுவினோடு

பசு + உடன் >பசுவினுடன்

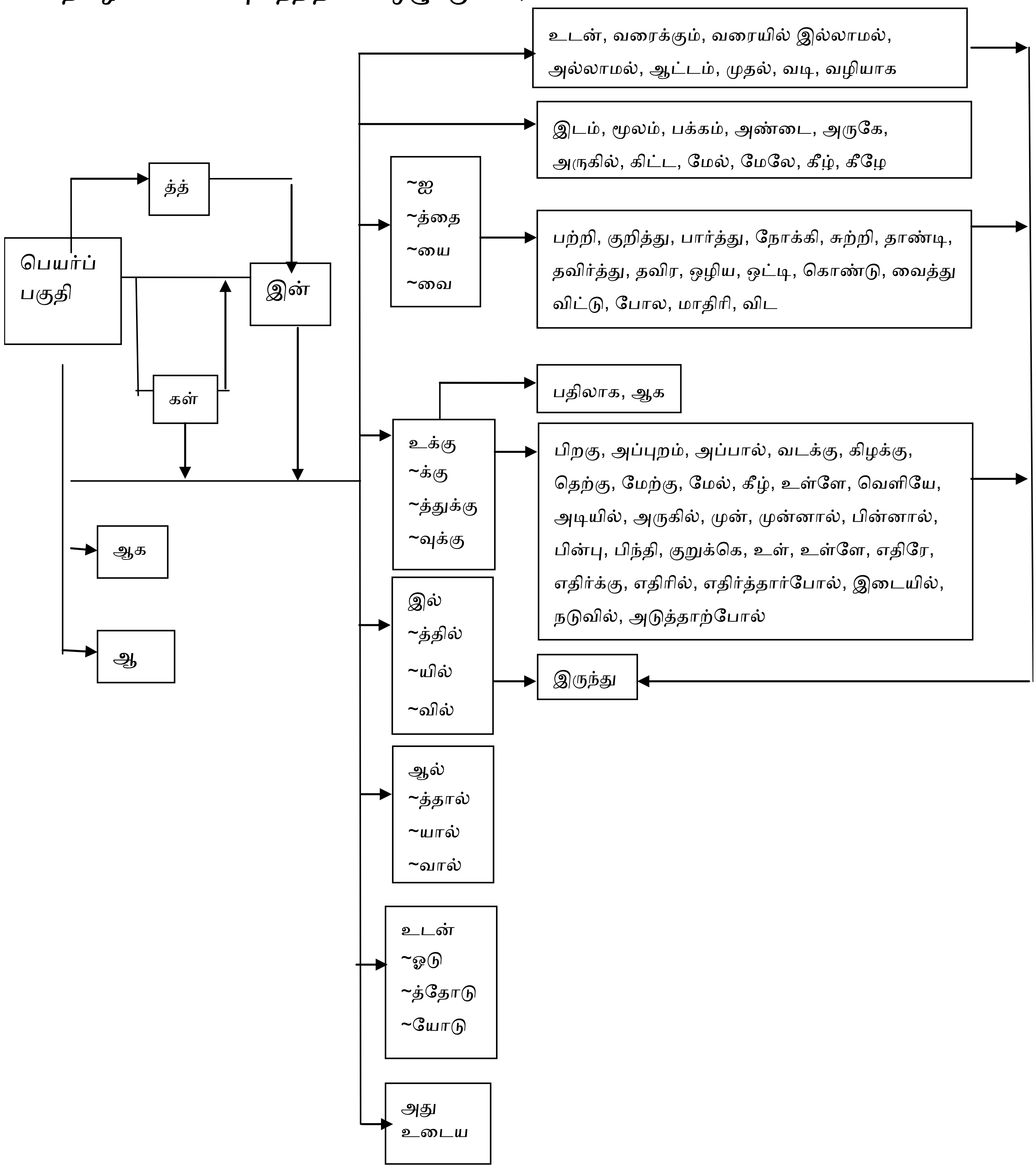
பசு + உடைய >பசுவினுடைய

பசு + அது >பசுவினது

பசு + இல் >பசுவினில்.

கணினி மூலம் பெயர்களை ஆய்வதற்காகக் கடா-வகுப்பு, எலி-வகுப்பு, ஈ-வகுப்பு, ஆறு-வகுப்பு, கண்-வகுப்பு, பொன்-வகுப்பு எனப் பெயர்கள் அவற்றின் உருபொலியனியல் நடத்தை அடிப்படையில் சில வகுப்புகளாகப் பிரிக்கப்பட வேண்டும் (பார்க்க இராஜேந்திரனின் உருபனியல் ஆய்வி உருவாக்கம் பற்றிய கட்டுரை).

தமிழ்ப் பெயர் வடிவத்திற்கான ஒழுக்கு வரைபடம்



3.3.2 வினைச்சொற்களின் உருபனியல் அமைப்பு

தமிழ் வினைகள் எல்லாம் கால இடைநிலைகளை ஏற்கும். அவைகள் வினைமுற்றுகளாகவோ வினையெச்சங்களாகவோ வரும். வினைமுற்றுகள் பின் வரும் அமைப்பைப் பெற்று வரும்:

வினையடி + கால ஒட்டு+ இடம்-எண்-பால் ஒட்டு

எ.கா.

நட-ந்த்-ஆன் > நடந்தான்

இதில் *நட* என்பது வினையடி, *ந்த்* என்பது இறந்தகாலம். இதைத் தொடர்ந்து வரும் *ஆன்* வினைமுற்று விகுதியாகும். கால ஒட்டுக்களை இறந்தகாலம், நிகழ்காலம், எதிர்காலம் என மூன்றாகப் பகுக்கலாம். எல்லாக் கால ஒட்டுக்களும் வினையடியை ஒட்டி அதனைத் தொடர்ந்து வரும்.

தமிழ் வினைகளைக் கால ஒட்டுகள், இட-எண்-பால் மற்றும் வினையின் எச்ச வடிவை அடையாளப்படுத்தும் வேறு சிலை ஒட்டுக்களை ஏற்று வரும் சொற்களாக வர்ணிக்கலாம். பெயர்களைப் போல் வினைகளும் குறையுள்ள உருபனியலைக் கொண்ட (எல்லா வினை ஒட்டுக்களையும் ஏற்காத) பல வினைகளைக் கொண்டிருக்கும்: அவை குறைவினைகள் (defective Verbs) எனப்படும்.

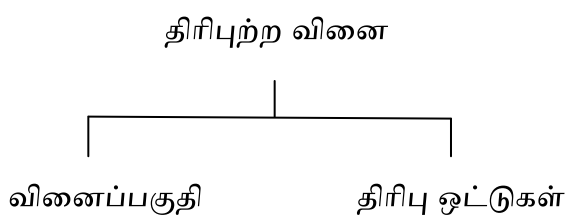
3.3.2.1 வினைத்திரிபு

தமிழ் வினைவடிவங்களை முற்று வடிங்கள் (finite form), எச்ச வடிவங்கள் (Non-finite forms), பெயராக்கம் செய்யப்பட்ட வடிவங்கள் (nominalized forms) என வேறுபடுத்தலாம். முற்று மற்றும் எச்ச வடிவங்களின் வர்ணனை உருபனியல் அடிப்படையிலானது என்று கூறலாம். நெகிழ்வாகச் சொன்னால் முற்றுவினை வடிவங்கள் முதன்மை எச்சத் தொடரின் (Main clause) பயனிலையாக வரும். அதாவது வாக்கியத்தை முற்றுப்பெறச் செய்யும். மாறாக எச்சவினை வடிவங்கள் உட்படு எச்சத்

தொடராகவோ துணைநிலை எச்சத்தொடராகவோ கூட்டுவினைக் கட்டமைப்புகளிலோ வரும்.

முற்றுவினை வடிவங்கள் வினைப்பாங்கு (mood) வகைப்பாடுகளுக்கு (அதாவது கட்டளை, வியங்கோள், உள்ளது கிளத்தல்) வேண்டி திரிபுறும். உள்ளது கிளத்தல் வினைகள் காலம், இடம், எண், பால் இவற்றின் வகைப்பாடுகளுக்காகத் திரிபுறும். எச்சவடிவங்கள் மற்றும் பெயராக்க வடிவங்கள் சில நேரங்களில் காலத்திற்கு வேண்டி திரிபுறும்; மேலும் வினையின் துணைநிலையை அல்லது பெயர்நிலையைக் குறிப்பிடும் சார்பு வகைப்பாட்டிற்கும் (relational cotregory) திரிபுறும். மேலும் முற்று மற்றும் எச்சவினை வடிவங்கள் (எல்லா நேர்வுகளிலும் இல்லாவிட்டாலும்) நேர்வடிவங்களையும் (positive forms) எதிர்மறை வடிவங்களையும் (Negative forms) கொண்டிருக்கும்.

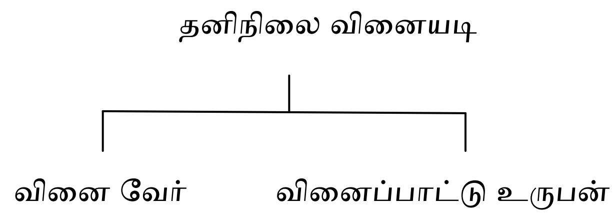
இவ்வாறு தமிழின் திரிபுற்ற வினை வடிவு (Inflected) பல உருபன்களின் ஒலியனியல் மெய்ப்படுத்தம் ஆகும் (எ.கா. வினை, கட்டளை, ஏவல், இடம், எண், பால் உருபன்கள், துணைநிலை ஆக்கம் செய்யும் உருபன்கள், எதிர்மறை இயக்கி). உருபன்களின் ஒலியனியல் மெய்படுத்தில் ஒரு உருபு (morph) ஒரு உருபனையோ (morpheme) பல உருபன்களையோ உருப்படுத்தம் செய்யும். எல்லா உருபுகளும் பகுதிவடிவில் தோன்றும் வினையுடன் சேர்க்கப்படும் திரிபுப் பின்னொட்டுகளாக வரும். அதைப் பின்வருமாறு உருப்படுத்தம் செய்யலாம்:



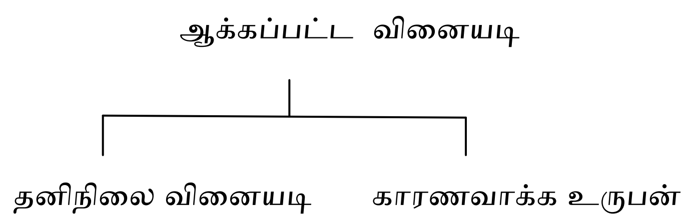
3.3.2.2. வினைப்பகுதி

வினைப்பகுதி என்பது திரிபு ஒட்டுக்கள் சேரக்கூடிய வினையின் வடிவமாகும். வினைப்பகுதி மேலும் முன்னிலை இடத்திற்கு ஏவல் ஒருமையாகப் பயன்படுத்தப்படும்

வினைவடிவமும் ஆகும். இவ்வடிவம் தான் அகராதியில் பதிவாகத் தரப்படும் வடிவமாகும். ஒரு வினையடி தனிநிலையானதாகவோ ஆக்கப்பட்டதாகவோ இருக்கலாம். ஒரு தனிநிலை அல்லது ஆக்கப்பட்ட வினையடி உருபனியல் அடிப்படையில் கலவைத் தன்மை உடையதாக இருக்கலாம். பின்னால் காட்டியுள்ளபடி ஒரு தனிநிலை வினைப்பகுதி வினைப்பாட்டு உருபன் (voice morpheme) சேர்க்கப்பட்டதாகவும் சொல் வேர் உருபனாகவும் (lexical root morpheme) இருக்கும்.



வினைப்பாட்டு உருபன் ஒன்றுகொன்று எதிராக நிற்கும் இரு உருபன்களில் ஒன்றைக் கொண்டிருக்கும். இந்த எதிர்நிலை உறுப்பினர்களில் ஒரு உறுப்பினர் இயக்கு வினைப்பாட்டையும் (affective voice) மற்றொன்று இயங்கு வினைப்பாட்டையும் (effective voice) வெளிப்படுத்தும் (Paramasivam, 1979). ஆக்கப்பட்ட வினைப்பகுதி தன்னிலை வினைப்பகுதியையும் காணவாக்கத்தைக் குறிக்கும் ஆக்கப் பின்னொட்டையும் கொண்டிருக்கும்.



3.3.2.3 வினைவேர்

வினைப்பகுதியின் முதலாவது மற்றும் கட்டாயமான அலகு வினைவேர் உருபனாகும். இது வினைப்பகுதியின் சொற்பொருண்மையைத் தருகிறது. வினைவேர் வினையின் குறைக்க இயலாச் சொல் உருபன் ஆகும்.

3.3.2.4 இயங்கு வினைப்பாட்டு மற்றும் இயக்கு வினைப்பாட்டு உருபன்

வினைவேர் ஒரு இரண்டாவது அலகால் தொடரப்படலாம். இது வினைப்பாட்டு உருபன் ஆகும். தமிழ் வினைகளில் ஏறக்குறைய 60 விழுக்காடு அவற்றின் பகுதியில் வினைப்பாட்டு உருபனைக் கொண்டிருக்கும் (Lehmann,1933:50). இம்மாதிரியான வினைப்பாட்டு உருபன் ஒரு வினைவேர் உருபனைத் தொடர்ந்து வந்தால், அவை விருப்பாக வராமல் எப்பொழுதும் கட்டாயமாக வரும். முன்னர் கூறியபடி வினைப்பாட்டு உருபன் ஒன்றுக்கொன்று எதிரிடையாக நிற்கும் இரண்டு உருபன்களில் ஒன்றைக் கொண்டிருக்கும். எதிர் இணைகளின் ஒரு உறுப்பினர் இயக்கு வினைப்பாட்டையும் மற்றொரு உறுப்பினர் இயங்கு வினைப்பாட்டையும் வெளிப்படுத்தும் (Paramasivam 1979). ஒரு வினைவேர் இந்த எதிரிணையின் ஒரு உறுப்பினருடன் சேர்ந்து வந்தால் அது மற்ற உறுப்பினருடனும் சேர்ந்து வரும். இருப்பினும் இரண்டு உருபன்களும் இணைந்து வராது. இவ்வாறு வினைப்பாடு உருபன்களுடன் வரும் வினை வேர்கள் இரு வினைப் பகுதிகளைக் கொண்டிருக்கும்; இவற்றில் ஒரு வினைப்பகுதி இயக்கு வினைப்பாட்டின் உருபனையும் மற்ற வினைப்பகுதி இயங்கு வினைப்பாட்டின் உருபனையும் கொண்டிருக்கும்.

எ.கா.

அவனுடைய தலை திரும்பியது – இயங்கு வினைப்பாடு

அவன் தலையைத் திருப்பினான் – இயக்கு வினைப்பாடு

திரும்பு என்ற இயங்கு வினைப்பகுதி, *திருப்பு* என்ற இயக்கு வினைப்பகுதிக்கும் *திரு* என்ற வேர் உருபன் அடிப்படையாகக் கொள்ளலாம். *ம்பு* என்பதை இயங்கு வினைப்பாட்டு மாற்றுருபாகவும், *ப்பு* என்பதை இயக்கு வினைப்பாட்டு மாற்றுருபாகவும் ஆயலாம். இவை தவிர வினைப்பாட்டு உருபன்களின் இணைகளை மெய்ப்படுத்தும் பிற மாற்றுருபுகளின் இணைகளும் இருக்கின்றன. சாதாரணமாக இயங்கு மாற்றுருபு தனி அடைப்பொலியையும் இயக்கு மாற்றுரு இரட்டித்த அடைப்பொலியையும் கொண்டிருக்கும்.

இயங்கு மாற்றுருபு	**இயக்கு மாற்றுருபு**	**எடுத்துக்காட்டு**
கு	க்கு	பெருகு, பெருக்கு
ந்து	த்து	திருந்து, திருத்து
ம்பு	ப்பு	நிரம்பு, நிரப்பு
டு	ட்டு	ஓடு, ஓட்டு

ஸ்டீவர் (Stever, 1983:80) என்பவர் இதற்கு மாறுபட்ட ஆய்வைத் தருகிறார். அவரால் *திரும்பு* அல்லது *திருப்பு* போன்ற வடிவம் இரு உருபன்களாக மெய்ப்படுத்தம் பெறும் ஒரு உருபாக எடுக்கப்படுகிறது: அதாவது வேர் உருபனாகவும் வினைப்பாட்டு உருபனாகவும் எடுக்கப்படுகிறது.

3.3.2.5 ஆக்கப்பட்ட வினைப்பகுதி

ஆக்கப்பட்ட வினைப்பகுதி தனி வினையடிகளில் இருந்து இயக்கு ஒட்டுகளான *வி*, *பி*, *ப்பி* என்பவைகளின் சேர்க்கையால் உருவாக்கப்படலாம். தற்காலத் தமிழில் இந்த ஆக்கச் செயற்பாங்கு மிகக் கட்டுப்பாட்டிற்கு உட்படுத்தப்பட்டுள்ளது.

எ.கா.

தெரி + வி = தெரிவி

அறி + வி = அறிவி

இயக்கு வினையாக்கப் பின்னொட்டு *வி* தான் பல வினையடிகளுடன் சேர்த்து வருகிறது. தனிவினைப் பகுதியிலிருந்து இயக்கு வினைப்பகுதியை ஆக்கமுடியுமா என்பதை வருவதுரைக்க இயலாது. எனவே இயக்கு வினைப்பகுதி அகராதியில் பட்டியலிடப்பட வேண்டும்.

3.3.2.6 மெல் வினைப்பகுதி, வல்வினைப்பகுதி, இடைவினைப்பகுதி

வினைப்பகுதி என்பது திரிபு ஒட்டுக்கள் சேர்க்க இயலும் முன்னிலை ஒருமை ஏவலாக வரும் மற்றும் அகராதியில் பட்டியலிடப்படும் வடிவமாக எடுத்துக் கொள்ளப்பட்டுகிறது. இதன்படி *தூங்கு*, *கேள்*, *நட* என்பன வினைப்பகுதிகளாகும். இவை ஏவல் வடிவில் வரும் இவற்றில் திரிபு பின்னொட்டுக்கள் சேர்க்கப்படும். இருப்பினும் பல வினையடிகள் (எ. கா. *கேள்*, *நட*) என்பன ஒரு ஒலியனாலோ ஒலியன் மயக்கத்தாலோ பெரிதாக்கப்படும். இக்கூடுதலான ஒலியனோ ஒலியன் மயக்கமோ வினையடியுடன் சேர்க்கப்படும். திரிபு ஒட்டு உயிர் ஒலியில் தொடங்கினால் தான் தோன்றும்; மெய்யொலியில் தொடங்கினால் தோன்றாது. இது சில வினைகளில் தனி அடைப்பொலியான *க்* ஆக வரும் (எ.கா. *கேள்* + *அ* > *கேட்க*); சிலவற்றுடன் இரட்டித்த அடைப்பொலியான *க்க்* ஆக வரும். (எ.கா. *நட* + *அ* > *நடக்க*). இந்த ஒலியனுக்கோ ஒலியன் மயக்கத்திற்கோ பொருண்மைச் செயல்பாடு இல்லை; ஒலியனில் செயல்பாடு இருக்கிறது. இதை ஒலியனியல் பின்னொட்டாகவோ வெற்று உருபனாகவோ (Empty Morph) ஆயலாம். உயிரொலியால் தொடங்கும் திரிபு பின்னொட்டைச் சேர்க்கும்போது இரட்டித்த அடைப்பொலி *க்க்*–ஆல் பெரிதாக்கப்படுகிற வினைப்பகுதி வல்வினைப்பகுதி எனப்படும். எடுத்துக்காட்டாக *நட* என்ற வினைப்பகுதி வல்வினைப்பகுதியாகும். இதில் உயிரொலியில் தொடங்கும் பின்னொட்டான *அ* என்ற எச்சப் பின்னொட்டு சேர்க்கப்படும் போது அது இரட்டித்த அடைப்பொலியான *க்க்* என்பதால் பெரிதாக்கப்படும் (எ.கா. *நட* + *அ* > *நடக்க*). மாறாக இறந்தகாலப் பின்னொட்டான *ந்த்* சேர்க்கப்பட்டால் வினைப்பகுதி *க்க்* ஆல் பெரிதாக்கப்படுவதில்லை (எ.கா. *நட* + *ந்த்* + *அ* > *நடந்த*).

தனி இடைப்பொலியான *க்* என்பதால் பெரிதாக்கப்படக்கூடிய வினைப்பகுதி நடு / இடை வினைப்பகுதியாகும். எடுத்துக்காட்டாக *கேள்* என்ற வினைப்பகுதி இடை வினைப்பகுதியாகும். உயிர் தொடக்கத் திரிபு பின்னொட்டான *அ* என்பதை இவ்வினையுடன் சேர்த்தால் வினைப்பகுதி தனி அடைப்பொலியான *க்* ஆல பெரிதாக்கப்படும் (எ.கா.*கேள்* + *அ* > *கேட்க*). மாறாக மெய்யொலி தொடக்கத் திரிபு ஒட்டு சேர்க்கப்பட்டால் வினைப்பகுதி பெரிதாக்கப்படாது (எ.கா. *கேள்* + *த்* + *ஏன்* > *கேட்டேன்*). உயிர் தொடக்கப் பின்னொட்டு சேர்க்கப்படும்போது பெரிதாகாத

வினைப்பகுதிகள் மெல்வினைப்பகுதி எனப்படும். எடுத்துக்காட்டாக *தூங்கு* என்பது மெல்வினைப்பகுதி ஆகும். இதில் *அ* என்ற உயிர் தொடக்க எச்சப் பின்னொட்டு சேர்க்கப்பட்டால் வினைப்பகுதி பெரிதாகாது (எ.கா. *தூங்கு* + *அ* > *தூங்க*).

தமிழ் வினைகளை எல்லாம் வல், இடை, மெல் வினைப்பகுதிக்கள் என்று பாகுபாடு செய்யலாம். பல்வேறுப்பட்ட திரிபுப் பின்னொட்டுக்களை விளக்குவதற்காகத் தமிழ் வினை வடிவங்களை முற்று வடிவங்கள் (finite forms). எச்ச வடிவங்கள் (non-finits forms) மற்றும் பெரயாக்கப்பட்ட வினை வடிவங்கள் என வேறுபடுத்தலாம்.

3.3.2.7 முற்று வினை வடிவங்கள்

முற்றுவினை வடிவங்களை வினைநோக்கு அடிப்படையில் பின்வருமாறு பிரிக்கலாம்.

1. ஏவல் வினை வடிவங்கள்
2. சுட்டு வினை வடிவங்கள்
3. வியங்கோள் வினை வடிவங்கள்

பிறவினை நோக்குகள் வினைத் திரிபுகளால் மெய்ப்படுத்தம் பெற்றுவதில்லை; துணை வினைகளால் மெய்ப்படுத்தம் பெறுகின்றன.

3.3.2.7.1. ஏவல் வினைவடிவங்கள்

தற்காலத் தமிழ் ஒருமை, பன்மை ஏவல் வடிவங்களை வேறுபடுத்துகிறது. ஒருமை, பன்மை வடிவங்கள் இரண்டு நேர்மறை, எதிர்மறை வடிவங்களாக வடிவமைக்கப்படும். ஏவல் வடிவங்களில் ஒரு வடிவம் இடத்தின் வகைப்பாட்டிற்காக வெளிப்படையாக அடையாளப்படுத்தப்படும். வேறுபட்ட ஏவல் வடிவங்களைப் பின்வருமாறு உள்ளுறை அமைப்பின் உருபன்களின் ஒலியனில் மெய்ப்படுத்தமாக ஆயலாம்:

வினை + ஏவல் + இடம் (+ எதிர்மறை), (+ பன்மை).

ஒருமை ஏவல் (நேர்மறை)

நேர்மறை ஒருமை ஏவல் வடிவம் வினைப்பகுதியை ஒத்திருக்கும். இதில் ஒட்டு இருக்காது.

எ.கா.

நட, கேள், தூங்கு

இது ஏவல் மற்றும் தன்மையிட உருபன் வெளிப்படையாக வெளிப்படுத்தப்படுவதில்லை என்றும் பூஜிய உருபனால் மெய்ப்படுத்தம் செய்யப்படும் என்றும் காட்டுகிறது.

ஒருமை ஏவல் (எதிர்மறை)

எதிர்மறை ஒருமை ஏவல் வடிவ *ஆத்* என்ற எதிர்மறை மாற்றுருபால் பின்னொட்டப்பட்ட வினைப்பகுதியைக் கொண்டிருக்கும். இது *ஏ* என்ற ஒலியிசைவுப் பின்னொட்டால் தொடரப்படும்.

வினைப்பகுதி + ஆத் + ஏ

எ.கா.

தூங்க + ஆத் + ஏ > தூங்காதே

கேட்க + ஆத் + ஏ > கேட்காதே

நடக் + ஆத் + ஏ > நடக்காதே

இந்த வடிவுகளின் ஏவல் மற்றும் முன்னிலை உருபன்கள் பூஜிய உருபனால் மெய்ப்படுத்தம் செய்யப்படும். எதிர்மறை உருபன் மட்டும் தான் *ஆத்* என்ற எதிர்மறை மாற்றுருபால் மெய்ப்படுத்தம் செய்யப்படும்.

பன்மை ஏவல் (நேர்மறை)

நேர்மறைப் பன்மை ஏவல் வடிவம் வினைப்பகுதி (உ)ங்கள் என்ற பன்மை ஒட்டையும் கொண்டிருக்கும்.

வினைப்பகுதி + (உ)ங்கள்

உங்கள் என்ற பன்மை ஒட்டை ஒருமை எண்ணுக்கும் உயர் நிலைக்கும், பன்மை எண்ணுக்கும் பொருள்கோள் செய்யலாம். இந்த ஏவல் வடிவை உயர்நிலையில் உள்ள தனிநபரைக் குறிப்பிடவோ ஒரு குழும தனிநபர்களைக் குறிப்பிடவோ பயன்படுத்தலாம்.

தூங்க்-உங்கள் > தூங்குங்கள்

நட-ங்கள் > நடங்கள்

கேள்-உங்கள் > கேளுங்கள்

ஏவல் மற்றும் பன்மை உருபன் (உ)*ங்கள்* என்ற உருபால் செய்ப்படுத்தம் செய்யப்படுவதாகவும் இந்த வடிவத்தில் உள்ள முன்னிலை இட உருபன் பூஜிய உருபனாக மெய்ப்படுத்தம் செய்யப்படுவதாகவும் ஆயலாம்.

பன்மை ஏவல் (எதிர்மறை)

எதிர்மறைப் பன்மை ஏவல் வடிவத்தில் வினைப்பகுதி *ஆத்* என்றஎதிர்மறை மாற்றுருபால் தொடரப்படும்: பின்னர் *ஈர்கள்* என்ற முன்னிலைப் பன்மைப் பின்னொட்டால் தொடரப்படும்.

வினைப்பகுதி +ஆத் + ஈர்கள்

எ.கா.

தூங்க் + ஆத் + ஈர்கள் > தூங்காதீர்கள்

கேட்க் + ஆத் + ஈர்கள் > கேட்காதீர்கள்

ஆத் என்ற மாற்றுருபால் எதிர்மறை ஏதிர்மறை உருபன் மெய்ப்படுத்தம் செய்யப்படும்போது முன்னிலை உருபனும் பன்மை உருபனும் *ஈர்கள்* என்ற ஒரே உருபால் மெய்ப்படுத்தம் செய்யப்படும். ஏவல் உருபன் பூஜிய உருபால் மெய்ப்படுத்தம் செய்யப்படும். நேர்மறைப் பன்மை ஏவல் வடிவு போல் மின்னிலைப் பன்மை வ்பின்னொட்டான ஈர்கள் என்பதை, ஒருமை, உயர்நிலை என்பனவற்றிற்காகவோ முன்னிலை மற்றும் பன்மைக்காகவோ பொருள்கோள் செய்யலாம்.

ஏவல் + பின்னொட்டுக்கள்

இரு வகையான பின்னொட்டுகள் ஏவல் வடிவங்களுடன் சேர்க்கப்படலாம்: இந்நேர்வில் ஒட்டுநிலையை அடைந்துள்ள *ஏன்* என்ற கேள்விச்சொல் மற்றும் பின்னொட்டுகளாக வந்து விளிக் குறிகளாகச் (address markers) செயல்படும் பலவைஅகிப்பட்ட பெயர்கள், குறிப்பாக உறவுமுறைச் சொற்கள். *ஏன்* என்ற வினாச் சொல் ஏவல் வடிவுடன் பின்னொட்டப்படும்போது ஏவலால் வெளிப்படும் கட்டளை அல்லது வேண்டல் ஒரு கருத்தாக மாற்றப்படும்.

எ.கா.

வாய் – ய் – ஏன் > வாயேன்

கேள் – உங்கள் – ஏன் > கேளுங்களேன்

அடா, *அடி*, *ஐயா*, *அப்பா*, *அம்மா*, *அண்ணன்*, *அக்கா* போன்ற பெயர்கள் பலவகைக் கட்டளை வடிவுகளில் சேரும்போது தொடக்க உயிரான *அ* என்பதை இழந்து பின்னொட்டு நிலையை அடைகின்றன.

எ.கா.

ஒருமை கட்டளை வடிவத்துடன்	**பன்மை கட்டளை வடிவத்துடன்**
வா – டா > வாடா	வார் – உங்க(ள்) – டா > வாருங்கடா
வா – டி > வாடி	வார் – உங்க(ள்) – டீ > வாருங்கடீ
வாய் – ய்யா > வாய்யா	வார் – உங்க(ள்) – ம்மா > வாருங்கம்மா
வா – ம்மா > வாம்மா	
வா – ப்பா > வாப்பா	
வா – ண்ணா > வாண்ணா	
வா – க்கா > வாக்கா	

3.3.2.7.2. சுட்டுவினை வடிவுகள்

தமிழில் நேரடியான மற்றும் எதிர்மறையான சுட்டுவினை வடிவங்கள் இருக்கின்றன. நேரிடைச் சுட்டுவினை வடிவங்கள் காலம், இடம், எண், பால் மற்றும்

நிலை இவற்றிற்காக அடையாளப்படுத்தப்பட்டிருக்கும். எதிர்மறைச் சுட்டுவினை வடிவங்கள் நேரிடைச் சுட்டுவினை வடிவங்களைப் போல் அடையாளப் படுத்தப்படுவதில்லை. நேரிடைச் சுட்டு வடிவங்களுக்கும் சுட்டு வடிவங்களுக்கும் ஒன்றுக்கொன்றான பொருத்தம் இல்லை.

நேரிடைச் சுட்டுவினை வடிவங்கள்

நேரிடைச் சுட்டுவினை வடிவங்கள் காலம், இடம், எண், பால், நிலை என்ற வகைப்பாடுகளுக்காக அடையாளப்படுத்தப்பட்டிருக்கும். இவ்வடிவங்கள் வினைப்பகுதி. கால உருபன் மற்றும் இடம், எண் மற்றும் பால் + நிலை இவற்றின் மூன்று உருபன்கள் என்பனவற்றைக் கொண்டு பின்வருமாறு உருப்படுத்தப்படும்:

வினைப்பகுதி + காலம் + இடம் + எண் + பால்

இந்த உருபன்களின் ஒலியனியல்சார் மெய்ப்படுத்தத்தில் கால உருபன் பல மாற்றுருபுகளால் மெய்ப்படுத்தம் செய்யப்படும். இதன் விளைவாகப் பெரும்பாலான நேரிடைச் சுட்டுவினை வடிவங்களை வினைப்பகுதியாகவும் இரு திரிபுப் பின்னொட்டுக்களாகவும் பகுக்கலாம்.

வினைப்பகுதி + கால பின்னொட்டு + எண் – இடம் – பால் பின்னொட்டு

3.3.2.7.3 வினைப்பகுதி

வினைப்பகுதி குறித்து முன்னர் கூறப்பட்டுள்ளது.

3.3.2.7.4. காலப் பின்னொட்டுக்கள்

தமிழ் இறந்தகாலம், நிகழ்காலம், எதிர்காலம் என்ற மூன்று காலங்களை வேறுப்படுத்துகிறது. ஒவ்வொரு கால உருபனும் பல காலப் பின்னொட்டுக்களால் அல்லது மாற்றுருபுகளால் மெய்ப்படுத்தம் செய்யப்படுகின்றன. நிகழ்காலத்துக்கு *கிற்*, *க்கிற்* என்ற மாற்றுருபுகள் இருக்கின்றன; *கின்று*, *க்கின்று* என்பன நடையியல் வேறுபாடுகளாகும். இறந்த காலத்துக்கு *த்*, *ந்த்*, *இன்*, *இ*, *த்த்* என்ற ஐந்து மாற்றுருபுகள் இருக்கின்றன. ஒரு குறிப்பிட்ட வினை ஒவ்வொரு குழுமத்தின் ஒரு மாற்றுருபுடன்

மட்டுமே வரும். இதற்கு விதிவிலக்காக வருபவை *ந்* மற்றும் *இ* என்ற இறந்த காலத்தின் மாற்றுருபுகள் அடிப்படியில் அதாவது ஒரு குறிப்பிட்ட வினைப்பகுதியுடன் வரும் கால மாற்றுருபு குழுமம் அடிப்படையில் தமிழ் வினைகளைப் பல வகுப்புகளாகப் பிரிக்கலாம். பல பாகுபாடுகள் இது தொடர்பான நூல்களில் தரப்பட்டுள்ளன. ஆர்டனில் (1942:144) தரப்பட்டுள்ள கிராலின் (Graul) பாகுபாடு மிகக் கூடுதலாக ஏற்றுக்கொள்ளப்பட்ட பாகுபாடாகும் (classification). இப்பாகுபாட்டின் படி தமிழ் வினைகள் எல்லாம் ஏழு வகுப்புகளாகப் பகுக்கப்பட்டுள்ளன. இவற்றில் இரண்டு வகுப்புகளுக்குத் துணை வகுப்புகள் இருக்கின்றன. பின்வரும் அட்டவணை ஏழு வகுப்புகளையும் அவற்றின் கால ஒட்டுக்களையும் தருகின்றது.

வகுப்பு	**நிகழ்காலம்**	**இறந்தகாலம்**	**எதிர்காலம்**
1.	- கிற் -	- த் -	- வ் -
2.	- கிற் -	- ந்த் -	- வ் -
3.	- கிற் -	இன்	- வ் -
4.	- கிற் -	இரட்டிப்பு	- வ் -
5.	- கிற் -	- த் –	- ப் -
6.	- க்கிற் -	- த்த் -	- ப்ப் -
7.	- க்கிற் -	- ந்த் -	- ப்ப் -

ஒன்றிலிருந்து நான்கு வகுப்புகளைச் சார்ந்த வினைகள் மெல் வினைகள் என்று ஐந்தாவது வகுப்பைச் சார்ந்த வினைகள் இடை வினைகள் என்றும் ஆறு, ஏழு வகுப்புகளைச் சார்ந்த வினைகள் வல் வினைகள் என்றும் அழைக்கப்படுகின்றன. நான்காவது வகுப்பைச் சார்ந்த வினைகள் *டு*, *று* அல்லது கு என்ற அசையால் இறுதியுறும்; இவ்வினைகளில் இறந்தகால உருபன் காலப் பின்னொட்டால் மெய்ப்படுத்தம் பெறுவதில்லை. மாறாக *ட்*, *ற்*, *க்* என்பன இரட்டுவதால் மெய்ப்படுத்தம் பெறும்.

எ.கா.

போடு > போட் - ட்

பெறு > பெற் - ற்

புகு > புக் – க்

இவ்வினை வகுப்புகளின் ஒலியனியல் செயல்பாடுகளும் திரிபு அடுக்குகளும் பின்னால் விளக்கப்பட்டுள்ளன.

3.3.2.7.5 இட-எண்-பால் பின்னொட்டுக்கள்

கால ஒட்டுக்களைத் தொடரும் இரண்டாவது திரிபு ஒட்டு. இட–எண்– பால் பின்னொட்டாகும். இதை மூன்று உருபன்களை மெய்ப்படுத்தம் செய்யும் போர்ட்மாண்டோ உருபாக (Portmanteaumorph) ஆயலாம். இப்பின்னொட்டால் எழுவாய்ப் பெயர்த்தொடர் முற்று வினையுடன் உடன்பாடு காட்டும். பின்வரும் அட்டவணையில் இந்த உடன்பாடு தரப்பட்டுள்ளது:

மாற்றுப்பெயர்கள்	இட – எண் – பால் பின்னொட்டுக்கள்
நான்	ஏன்
நீ	ஆய்
அவன்	ஆன்
அவள்	ஆள்
அது	அது, து, உம்
அவர்	ஆர்
நாங்கள், நாம்	ஓம்
நீங்கள்	ஈர்கள்
அவர்கள்	ஆர்கள்
அவை	அ

3.3.2.7.6 திரிபு அடுக்குகள்

கால் ஒட்டுக்கள் மற்றும் இட-எண்-பால் ஒட்டுக்கள் கொண்ட இனம் சுட்டு வினையின் (indicative verb) திரிபு நேரடியானது. ஒரு குறிப்பிட்ட வகுப்பைச் சார்ந்த வினையடி அந்த வகுப்பிற்கென்று அனுமதிக்கப்பட்ட கால மாற்றுருபுகளின் குழுமத்தை

ஏற்கும். அடைத் தொடர்ந்து இடம்-எண்-பால் ஒட்டு சேர்க்கப்படும். காலப் பின்னொட்டுக்கள் மற்றும் இடம்-எண்-பால் பின்னொட்டுக்களின் வருகை முறையில் திரிபு அடுக்கில் சீரற்ற வினை வடிவங்களும் இருக்கின்றன. கீழ்வரும் அட்டவணைகளில் ஒவ்வொரு வினை வகுப்புகளையும் சார்ந்த வினைகளின் திரிபு அடுக்குகள் தரப்பட்டுள்ளன.

வகுப்பு 1

இடம்	**நிகழ்காலம்**	**இறந்தகாலம்**	**எதிர்காலம்**
தன்மை ஒருமை	அழு-கிற்-ஏன்	அழு-த்-ஏன்	அழு-வ்-ஏன்
முன்னிலை ஒருமை	அழு-கிற்-ஆய்	அழு-த்-ஆய்	அழு-வ்-ஆய்
படர்க்கை ஒருமை ஆண்பால்	அழு-கிற்-ஆன்	அழு-த்-ஆன்	அழு-வ்-ஆன்
படர்க்கை ஒருமை பெண்பால்	அழு-கிற்-ஆள்	அழு-த்-ஆள்	அழு-வ்-ஆள்
படர்க்கை ஒருமை அஃறினை	அழு-கிற்-அது	அழு-த்-அது	அழு-ம்
படர்க்கை ஒருமை ஆண்பால் - பெண்பால்	அழு-கிற்-ஆர்	அழு-த்-ஆர்	அழு-வ்-ஆர்
தன்மைப்பன்மை	அழு-கிற்-ஓம்	அழு-த்- ஓம்	அழு-வ்- ஓம்
முன்னிலைப்பன்மை	அழு-கிற்-ஈர்கள்	அழு-த்- ஈர்கள்	அழு-வ்-ஈர்கள்
படர்க்கை பன்மை ஆண்பால் – பெண்பால்	அழு-கிற்-ஆர்கள்	அழு-த்-ஆர்கள்	அழு-வ்-ஆர்கள்
படர்க்கை பன்மை அஃறிணை	அழு-கின்ற்-அன்	அழு-த்-அன் -அ	அழு-ம்

வகுப்பு2

இடம்	நிகழ்காலம்	இறந்தகாலம்	எதிர்காலம்
தன்மை ஒருமை	விழு-கிற்-ஏன்	விழு-ந்த்-ஏன்	விழு-வ்-ஏன்
முன்னிலை ஒருமை	விழு-கிற்-ஆய்	விழு-ந்த்-ஆய்	விழு-வ்-ஆய்
படர்க்கை ஒருமை ஆண்பால்	விழு-கிற்-ஆன்	விழு-ந்த்-ஆன்	விழு-வ்-ஆன்
படர்க்கை ஒருமை பெண்பால்	விழு-கிற்-ஆள்	விழு-ந்த்-ஆள்	விழு-வ்-ஆள்
படர்க்கை ஒருமை அஃறினை	விழு-கிற்-அது	விழு-ந்த்-அது	விழு-ம்
படர்க்கை ஒருமை ஆண்பால் - பெண்பால்	விழு-கிற்-ஆர்	விழு-ந்த்-ஆர்	விழு-வ்-ஆர்
தன்மைப்பன்மை	விழு-கிற்-ஓம்	விழு-ந்த்- ஓம்	விழு-வ்- ஓம்
முன்னிலைப்பன்மை	விழு-கிற்-ஈர்கள்	விழு-ந்த்- ஈர்கள்	விழு-வ்- ஈர்கள்
படர்க்கை பன்மை ஆண்பால் - பெண்பால்	விழு-கிற்-ஆர்கள்	விழு-ந்த்-ஆர்கள்	விழு-வ்-ஆர்கள்
படர்க்கை பன்மை அஃறிணை	விழு-கின்ற்-அ	விழு-ந்த்-அன் -அ	விழு-ம்

வகுப்பு3

இடம்	**நிகழ்காலம்**	**இறந்தகாலம்**	**எதிர்காலம்**
தன்மை ஒருமை	தூங்கு-கிற்-ஏன்	தூங்கு-இன்-ஏன்	தூங்கு-வ்-ஏன்
முன்னிலை ஒருமை	தூங்கு-கிற்-ஆய்	தூங்கு-இன்-ஆய்	தூங்கு-வ்-ஆய்
படர்க்கை ஒருமை ஆண்பால்	தூங்கு-கிற்-ஆன்	தூங்கு-இன்-ஆன்	தூங்கு-வ்-ஆன்
படர்க்கை ஒருமை பெண்பால்	தூங்கு-கிற்-ஆள்	தூங்கு-இன்-ஆள்	தூங்கு-வ்-ஆள்
படர்க்கை ஒருமை அஃறினை	தூங்கு-கிற்-அது	தூங்கு-இன்-அது	தூங்கு-ம்
படர்க்கை ஒருமை ஆண்பால் - பெண்பால்	தூங்கு-கிற்-ஆர்	தூங்கு-இன்-ஆர்	தூங்கு-வ்-ஆர்
தன்மைப்பன்மை	தூங்கு-கிற்-ஓம்	தூங்கு-இன்- ஓம்	தூங்கு-வ்- ஓம்
முன்னிலைப்பன்மை	தூங்கு-கிற்-ஈர்கள்	தூங்கு-இன்- ஈர்கள்	தூங்கு-வ்- ஈர்கள்
படர்க்கை பன்மை ஆண்பால் - பெண்பால்	தூங்கு-கிற்-ஆர்கள்	தூங்கு-இன்-ஆர்கள்	தூங்கு-வ்-ஆர்கள்
படர்க்கை பன்மை அஃறிணை	தூங்கு-கிற்-அன்-அ	தூங்கு-இன்-அ	தூங்கு-ம்

வகுப்பு4

இடம்	**நிகழ்காலம்**	**இறந்தகாலம்**	**எதிர்காலம்**
தன்மை ஒருமை	போடு-கிற்-ஏன்	போட்-ட்-ஏன்	போடு- -வ்-ஏன்
முன்னிலை ஒருமை	போடு-கிற்-ஆய்	போட்-ட்-ஆய்	போடு-வ்-ஆய்
படர்க்கை ஒருமை ஆண்பால்	போடு-கிற்-ஆன்	போட்-ட்-ஆன்	போடு-வ்-ஆன்
படர்க்கை ஒருமை பெண்பால்	போடு-கிற்-ஆள்	போட்-ட்-ஆள்	போடு-வ்-ஆள்
படர்க்கை ஒருமை அஃறினை	போடு-கிற்-அது	போட்-ட்-அது	போட்-உம்
படர்க்கை ஒருமை ஆண்பால்-பெண்பால்	போடு-கிற்-ஆர்	போட்-ட்-ஆர்	போடு-வ்-ஆர்
தன்மைப்பன்மை	போடு-கிற்-ஓம்	போட்-ட்- ஓம்	போடு-வ்- ஓம்
முன்னிலைப்பன்மை	போடு-கிற்-ஈர்கள்	போட்-ட்- ஈர்கள்	போடு-வ்- ஈர்கள்
படர்க்கை பன்மை ஆண்பால் - பெண்பால்	போடு-கிற்-ஆர்கள்	போட்-ட்-ஆர்கள்	போடு-வ்-ஆர்கள்
படர்க்கை பன்மை அஃறிணை	போடு-கின்ற-அன்-அ	போட்-ட்-அன்-அ	போட்-உம்

வகுப்பு5

இடம்	**நிகழ்காலம்**	**இறந்தகாலம்**	**எதிர்காலம்**
தன்மை ஒருமை	கேட்-கிற்-ஏன்	கேட்-ட்-ஏன்	கேட்-ப்-ஏன்
முன்னிலை ஒருமை	கேட்-கிற்-ஆய்	கேட்-ட்-ஆய்	கேட்-ப்-ஆய்
படர்க்கை ஒருமை ஆண்பால்	கேட்-கிற்-ஆன்	கேட்-ட்-ஆன்	கேட்-ப்-ஆன்
படர்க்கை ஒருமை பெண்பால்	கேட்-கிற்-ஆள்	கேட்-ட்-ஆள்	கேட்-ப்-ஆள்
படர்க்கை ஒருமை அஃறினை	கேட்-கிற்-அது	கேட்-ட்-அது	கேட்-க்-உம்
படர்க்கை ஒருமை ஆண்பால்-பெண்பால்	கேட்-கிற்-ஆர்	கேட்-ட்-ஆர்	கேட்-ப்-ஆர்
தன்மைப்பன்மை	கேட்-கிற்-ஓம்	கேட்-ட்- ஓம்	கேட்-ப்- ஓம்
முன்னிலைப்பன்மை	கேட்-கிற்-ஈர்கள்	கேட்-ட்- ஈர்கள்	கேட்-ப்- ஈர்கள்
படர்க்கை பன்மை ஆண்பால் - பெண்பால்	கேட்-கிற்-ஆர்கள்	கேட்-ட்-ஆர்கள்	கேட்-ப்-ஆர்கள்
படர்க்கை பன்மை அஃறிணை	கேட்-கிற்-அன்-அ	கேட்-ட்-அன்-அ	கேட்-க்-உம்

வகுப்பு6

இடம்	**நிகழ்காலம்**	**இறந்தகாலம்**	**எதிர்காலம்**
தன்மை ஒருமை	படி-க்கிற்-ஏன்	படி-த்த்-ஏன்	படி-ப்ப்-ஏன்
முன்னிலை ஒருமை	படி-க்கிற்-ஆய்	படி-த்த்-ஆய்	படி-ப்ப்-ஆய்
படர்க்கை ஒருமை ஆண்பால்	படி-க்கிற்-ஆன்	படி-த்த்-ஆன்	படி-ப்ப்-ஆன்

படர்க்கை ஒருமை பெண்பால்	படி-க்கிற்-ஆள்	படி-த்த்-ஆள்	படி-ப்ப்-ஆள்
படர்க்கை ஒருமை அஃறினை	படி-க்கிற்-அது	படி-த்த்-அது	படி-க்க்-உம்
படர்க்கை ஒருமை ஆண்பால் - பெண்பால்	படி-க்கிற்-ஆர்	படி-த்த்-ஆர்	படி-ப்ப்-ஆர்கள்
தன்மைப்பன்மை	படி-க்கிற்-ஓம்	படி-த்த்- ஓம்	படி-ப்ப்- ஓம்
முன்னிலைப்பன்மை	படி-க்கிற்-ஈர்கள்	படி-த்த்- ஈர்கள்	படி-ப்ப்- ஈர்கள்
படர்க்கை பன்மை ஆண்பால் - பெண்பால்	படி-க்கிற்-ஆர்கள்	படி-த்த்-ஆர்கள்	படி-ப்ப்-ஆர்கள்
படர்க்கை பன்மை அஃறிணை	படி-க்கின்ற-அன்-அ	படி-த்த்-அன்-அ	கேட்-க்க்-உம்

வகுப்பு7

இடம்	**நிகழ்காலம்**	**இறந்தகாலம்**	**எதிர்காலம்**
தன்மை ஒருமை	நட-க்கிற்-ஏன்	நட-ந்த்-ஏன்	நட-ப்ப்-ஏன்
முன்னிலை ஒருமை	நட-க்கிற்-ஆய்	நட-ந்த்-ஆய்	நட-ப்ப்-ஆய்
படர்க்கை ஒருமை ஆண்பால்	நட-க்கிற்-ஆன்	நட-ந்த்-ஆன்	நட-ப்ப்-ஆன்
படர்க்கை ஒருமை பெண்பால்	நட-க்கிற்-ஆள்	நட-ந்த்-ஆள்	நட-ப்ப்-ஆள்
படர்க்கை ஒருமை அஃறினை	நட-க்கிற்-அது	நட-ந்த்-அது	நட-க்க்-உம்
படர்க்கை ஒருமை	நட-க்கிற்-ஆர்	நட-ந்த்-ஆர்	நட-ப்ப்-ஆர்கள்

ஆண்பால் - பெண்பால்			
தன்மைப்பன்மை	நட-க்கிற்-ஓம்	நட-ந்த்- ஓம்	நட-ப்ப்- ஓம்
முன்னிலைப்பன்மை	நட-க்கிற்-ஈர்கள்	நட-ந்த்- ஈர்கள்	நட-ப்ப்- ஈர்கள்
படர்க்கை பன்மை ஆண்பால் - பெண்பால்	நட-க்கிற்-ஆர்கள்	நட-ந்த்-ஆர்கள்	நட-ப்ப்-ஆர்கள்
படர்க்கை பன்மை அஃறிணை	நட-க்கின்ற-அன்-அ	நட-ந்த்-அன்-அ	நட-க்க்-உம்

3.3.2.7.7. வினைகளின் சீரற்றவடிவங்கள்

மேற்சொன்ன ஏழு வகுப்புகளில் அடக்க இயலாத ஒரு சிறு குழும வினைகள் இருக்கின்றன. இவ்வினைகள் பின்வரும் காரணங்களால் சீரற்ற வடிவங்களாக இருக்கின்றன.

1. அவை சீரற்ற கால ஒட்டை ஏற்கும். எடுத்துக்காட்டாகச் *சொல்* என்ற வினை ன் என்ற இறந்தகால ஒட்டை ஏற்கும் (*சொ – ன் – ஏன் > சொன்னேன்*).

2. திரிபு ஒட்டுக்கள் சேர்க்கப்படும்போது அவை தமது ஒலியன் வடிவை மாற்றுகின்றன. எடுத்துக்காட்டாகச் *சா* என்ற வினை இறந்தகாலத்திற்காகத் திரிபுறும்போது அதன் வினைப்பகுதி *செ* என்று மாறும் (*செ-த்த்-ஏன் > செத்தேன்*).

3.3.2.7.8 வியங்கோள் வடிவம்

வியங்கோள் வடிவம் வியங்கோள் பின்னொட்டான *க* வினைப்பகுதியுடன் சேர்க்கப்படுவதால் உருவாக்கப்படுகிறது. இப்பின்னொட்டு 6-ஆவது மற்றும் 7-ஆவது வகுப்பு வினைகளுடன் *க்க்* என்ற மாற்றுருபாக உருப்படுத்தம் பெறும்.

எ.கா

வாழ் – க > வாழ்க

வீழ் – க > வீழ்க

நட – க்க > நடக்க

கொடு – க்க > கொடுக்க

3.3.2.8 முற்றல்லா வினை வடிவங்கள்

நான்கு முற்றல்லா வினை வடிவங்களை வேறுபடுத்தலாம்:

1. செய்ய வினையெச்சம்
2. செய்து வினையெச்சம்
3. நிபந்தனை எச்சம்
4. பெயரெச்சம்

3.3.2.8.1 செய்ய வினையெச்சம்

இது வினையெச்சப் பின்னொட்டான *அ* வினைப்பகுதியுடன் சேர்க்கப்படுவதால் உருவாக்கப்படுகின்றது. இப்பின்னொட்டு மூன்று மாற்றுருபுகளைக் கொண்ட ஒரு உருபன் ஆகும். *அ* என்ற பின்னொட்டு ஒன்றாம், இரண்டாம், மூன்றாம், நான்காம் வகுப்பு வினைகளுடனும், *க* என்ற பின்னொட்டு ஐந்தாவது வகுப்பு வினைகளுடனும் *க்க* என்ற பின்னொட்டு ஆறாவது மற்றும் ஏழாவது வகுப்பு வினைகளுடன் வரும்.

அழு (வகுப்பு 1) : அழ் – அ

உட்கார் (வகுப்பு 2) : உட்கார் – அ

தூங்கு (வகுப்பு 3) :தூங்க் – அ

போடு (வகுப்பு 4) : போட் – அ

கேள் (வகுப்பு 5) : கேட் – க

படி (வகுப்பு 6) : படி – க்க

நட (வகுப்பு 7) : நட – க்க

3.3.2.8. 2 செய்து / செய்யாமல் வினையெச்ச வடிவம்

செய்து வினையெச்சத்திற்கு நேர்மறை மற்றும் எதிர்மறை என்ற இரண்டு வடிவங்கள் இருக்கின்றன. நேர்மறை வினையெச்ச வடிவம் வினைப்பகுதியுடன் வினையெச்ச ஒட்டைச் சேர்ப்பதால் உருவாக்கப்படும். இவ்வினையெச்ச ஒட்டுக்கள் இறந்தகாலப் பின்னொட்டுக்களின் ஒலியன் வடிவைக் கொண்டிருக்கும். இவை தவிர பிற பின்னொட்டுக்கள் உ என்ற ஒலித்துணை உயிரைக் (Enunciative vowel) கொண்டிருக்கும். இவ்வாறு *து*, *ந்து*, *த்து*, *இ* என்ற நான்கு ஒட்டுக்களும் நேர் வினையெச்ச ஒட்டுக்களாக வரும்.

அழு (வகுப்பு 1) : அழு – து

உட்கார் (வகுப்பு 2) : உட்கார் – ந்து

தூங்கு (வகுப்பு 3) : தூங்க் – இ

போடு (வகுப்பு 4) : போட் – டு

கேள் (வகுப்பு 5) : கேட் – டு

படி (வகுப்பு 6) : படி – த்து

நட (வகுப்பு 7) : நட – ந்து

எதிர்மறையைப் பொருத்தவரையில் இரண்டு எதிர்மறை வினையெச்ச வடிவங்கள் இருக்கின்றன. முதல் வடிவில் *ஆ* என்ற எதிர்மறை மாற்றுருபு வினைப்பகுதியுடன் முதலில் சேர்க்கப்பட்டு அதன் பின் *மல்* என்ற வினையெச்சப் பின்னொட்டு சேர்க்கப்படும். இவ்வாறு முதல் எதிர்மறை பெயரெச்ச வடிவத்தின் அமைப்பு புன்வருமாறு அமையும்.

வினைப்பகுதி + ஆ – மல்

இரண்டாவது வடிவத்தில் வினைப்பகுதியைத் தொடர்ந்து *ஆத்* என்ற எதிர்மறை மாற்றுருபு சேர்க்கப்பட்டு அதைத் தொடர்ந்து *உ* என்ற ஒலித்துணை உயிரொலி சேர்க்கப்படும். இவ்வாறு இரண்டாவது எதிர்மறை வினையெச்ச வடிவம் பின்வருமாறு அமையும்:

வினைப்பகுதி + ஆத் – உ

எ.கா.

அழு (வகுப்பு 1)	: ஆழ் – ஆ – மல்	ஆழ் – ஆது
உட்கார் (வகுப்பு 2)	: உட்கார் – ஆ – மல்	உட்கார் – ஆது
தூங்கு (வகுப்பு 3)	: தூங்க் – ஆ – மல்	தூங்க் – ஆது
போடு (வகுப்பு 4)	: போட் – ஆ – மல்	போட் – ஆது
கேட் (வகுப்பு 5)	:கேட் – க் – ஆ – மல்	கேட் - க் – ஆது
படி (வகுப்பு 6)	: படி – க்க - ஆ – மல்	படி – க்க – ஆது
நட (வகுப்பு7)	: நட –க்க – ஆ – மல்	நட – க்க – ஆது

3.3.2.8.3. நிபந்தனை வடிவங்கள்

மூன்றாவது காலம் காட்டாத முற்றல்லா வடிவம் நிபந்தனை வடிவம் ஆகும். இது எதிர்மறை வடிவத்திலும் நேர்மறை வடிவத்திலும் வரும். நேர்மறை வடிவம் வினைப்பகுதியுடன் பொருத்தமுடைய இறந்தகால மாற்றுருபை ஒக்கும். ஒலியன் கூறையும் அதைத் தொடர்ந்து *ஆல்* என்ற நிபந்தனைப் பின்னொட்டையும் பெற்றுவரும். நிபந்தனை வினை வடிவத்தின் அமைப்பு பின்வருமாறு அமையும்:

வினைப்பகுதி + பொருத்தமான இறந்தகால ஒட்டை ஒக்கும் ஒலியன் கூறு + ஆல்

இறந்தகால ஒட்டை ஒக்கும் ஒலியன் கூறு நிகழ்கால மாற்றுருபுடனோ எதிர்கால மாற்றுருபுடனோ முரண்படாது. இதற்குப் பொருண்மை செயற்பாடு இல்லை. அதாவது இது இறந்தகாலத்தை உணர்த்தாது.

அழு (வகுப்பு 1) : ஆழ் – த் – ஆல்

உட்கார் (வகுப்பு 2) : உட்கார் – ந்த்– ஆல்

தூங்கு (வகுப்பு 3) : தூங்க் – இன் – ஆல்

போடு (வகுப்பு 4) : போட் – ட் – ஆல்

கேட் (வகுப்பு 5) : கேட் – ட் – ஆல்

படி (வகுப்பு 6) : படி – த்த் – ஆல்

நட (வகுப்பு7) : நட –ந்த் – ஆல்

எதிர்மறை நிபந்தனை வினை வடிவம் வினைப்பகுதியுடன் ஆ என்ற எதிர்மறை உருபையும் அதைத் தொடர்ந்து *விட்டால்* என்ற நிபந்தனை ஒட்டையும் சேர்க்கப்படும் போது பெறப்படும்.

விழு (வகுப்பு 1) : விழ் – ஆ – விட்டால்

உட்கார் (வகுப்பு 2) : உட்கார் – ஆ – விட்டால்

தூங்கு (வகுப்பு 3) : தூங்க் – ஆ – விட்டால்

போடு (வகுப்பு 4) : போட் – ஆ – விட்டால்

கேட் (வகுப்பு 5) :கேட் – க் – ஆ – விட்டால்

படி (வகுப்பு 6) : படி – க்க - ஆ – விட்டால்

நட (வகுப்பு7) : நட –க்க – ஆ – விட்டால்

3.3.2.8.4. பெயரெச்ச வடிவங்கள்

பெயரெச்ச வடிவம் காலத்தை வேறுபடுத்தும் எச்ச வடிவமாகும். பெயரெச்ச வடிவம் இறந்தகாலப் பெயரெச்சம், நிகழ்காலப் பெயரெச்சம், எதிர்கால பெயரெச்சம் எனப்படும். இவற்றுடன் காலங்காட்டாத எதிர்மறைப் பெயரெச்சமும் சேரும். இறந்தகாலப் பெயரெச்ச வடிவமும் நிகழ்காலப் பெயரெச்ச வடிவமும் வினைப்பகுதியுடன்

மொத்தமான இறந்தகால மாற்றுருபையோ நிகழ்கால மாற்றுருபையோ சேர்த்தும் அதைத் தொடர்ந்து பெயரெச்ச பின்னொட்டான *அ* என்பதை சேர்த்தும் உருவாக்கப்படும்.

அழு (வகுப்பு 1)	: அழு – த் – அ	அழு – கிற் - அ
உட்கார் (வகுப்பு 2)	: உட்கார் – ந்த்– அ	உட்கார் – கிற் - அ
தூங்கு (வகுப்பு 3)	: தூங்க் – இ – அ	தூங்கு – கிற் - அ
போடு (வகுப்பு 4)	: போட் – ட் – அ	போடு – கிற் - அ
கேட் (வகுப்பு 5)	: கேட் – ட் – அ	கேட் – கிற் - அ
படி (வகுப்பு 6)	: படி – த்த் –அ	படி – க்கிற் - அ
நட (வகுப்பு7)	: நட –ந்த் – அ	நா – க்கிற் – அ

எதிர்கால பெயரெச்ச வடிவம் *உம்* என்ற ஒட்டை வினைப்பகுதியுடன் சேர்ப்பதால் பெறப்படும். *உம்* என்ற என்ற ஒரு போர்ட்மாண்டோ உருபாகும்; இது எதிர்கால உருபனையும் பெயரெச்ச உருபனையும் மெய்ப்படுத்தம் செய்யும்.

அழு (வகுப்பு 1)	: அழ் – உம்
உட்கார் (வகுப்பு 2)	: உட்கார் – உம்
தூங்கு (வகுப்பு 3)	: தூங்க் – உம்
போடு (வகுப்பு 4)	: போட் – உம்
கேட் (வகுப்பு 5)	:கேட் – க் – உம்
படி (வகுப்பு 6)	: படி – க் - உம்
நட (வகுப்பு7)	: நட –க் – உம்

எதிர்மறைப் பெயரெச்ச வடிவம் எதிர்மறை பெயரான *ஆத்* என்பதையும் அதைத் தொடர்ந்து பெயரெச்ச ஒட்டான *அ* என்பதையும் வினைப்பகுதியுடன் சேர்த்துப் பெறப்படும்.

அழு (வகுப்பு 1) : அழ் – ஆத் - அ

உட்கார் (வகுப்பு 2) : உட்கார் – ஆத் - அ

தூங்கு (வகுப்பு 3) : தூங்க் – ஆத் - அ

போடு (வகுப்பு 4) : போட் – ஆத் - அ

கேள் (வகுப்பு 5) :கேட் – க் – ஆத் - அ

படி (வகுப்பு 6) : படி – க்க் - ஆத் - அ

நட (வகுப்பு7) : நட –க்க் – ஆத் - அ

3.3.2.9 பெயராக்கப்பட்ட வினை வடிவங்கள் / தொழிற்பெயர்கள்

தொழிற்பெயர்களைக் காலக்குறி பெற்றவை, காலக்குறி பெறாதவை என இரண்டு வகைகளாக வேறுபடுத்தலாம். காலக்குறி பெறாத தொழிற்பெயர்கள் வினைப்பகுதியும் அதைத் தொடர்ந்து பெயராக்க ஒட்டையும் பெற்றிருக்கும். இவ்வகையிற்படும் தொழிற்பெயர்கள் மூன்று வேறுபட்ட பெயராக்க ஒட்டுக்களைப் பெற்று வரும்:

1. வினைப்பகுதி + ஆல்

செய் (வகுப்பு 1) : செய் – ய் – அல்

படி (வகுப்பு 6) : படி – க்க் – அல்

2. வினைப்பகுதி + தல்

செய் (வகுப்பு 1) : செய் – தல்

படி (வகுப்பு 6) : படி - த்தல்

3. வினைப்பகுதி + கை

செய் (வகுப்பு 1) : செய் – கை

படி (வகுப்பு 6) : படி – க்கை

காலக்குறி பெற்ற தொழிற்பெயர்களைப் பொறுத்தவரையில் அவை வினைப்பகுதியுடன் மூன்று கால மாற்றுருபுகளையும் அதைத்தொடர்ந்து *அது* என்ற பெயராக்க ஒட்டையும்

பெற்று வரும். எதிர்மறைக்குறி பெற்ற தொழிற்பெயர்கள் *ஆத்* என்ற எதிர்மறை ஒட்டையும் அதைத் தொடர்ந்து *அது* என்ற பெயராக்க ஒட்டையும் பெற்று வரும்.

வினைப்பகுதி + இறந்தகால / நிகழ்கால / எதிர்கால / ஒட்டு + பெயராக்க ஒட்டு – அது

எ.கா.

செய் (வகுப்பு 1) : செய் – த் – அது
செய் – கிற் – அது
செய் – வ் – அது
செய் – ய் – ஆத் – அது

படி (வகுப்பு 6) : படி – த்த் – அது
படி – க்கிற் – அது
படி – ப்ப் – அது
படி – க்க் – ஆத் – அது

3.3.2.9.1 வினையாலணையும் பெயர்கள் மற்றும் பெயரடைப் பெயர்கள்

வினையாலணையும் பெயர்களைப் பின்வருமாறு ஆயலாம்:

வினைப்பகுதி + கால ஒட்டு / எதிர்மறை ஒட்டு + படர்க்கை இடச்சேய்மைச் சுட்டு மாற்றுப்பெயர்கள்.

எ.கா.

செய் – த் – அவன்	செய் – ப் – அவன்
செய் – த் – அவள்	செய் – ப் – அவள்
செய் – த் – அது	செய் – ப் – அது
செய் – த் – அவர்	செய் – ப் – அவர்
செய் – த் – அவர்கள்	செய் – ப் – அவர்கள்
செய் – த் – அவை	செய் – ப் – அவை
செய் – கிற் – அவன்	செய் – ய் - ஆத் – அவன்

செய் – கிற் – அவள் செய் – ய் - ஆத் – அவள்

செய் – கிற்– அது செய் – ய் - ஆத் – அது

செய் – கிற் – அவர் செய் – ய் - ஆத் – அவர்

செய் – கிற் – அவர்கள் செய் – ய் - ஆத் – அவர்கள்

செய் – கிற் – அவை செய் – ய் - ஆத் – அவை

3.3.2.10 குறைவினைகள்

சில வினைகள் எல்லாக் கால வகுப்புகளையும் எல்லா இடம்-எண்-பால் ஒட்டுக்களையும் முற்றல்லா எச்ச வடிவங்களையும் பெறாது வரும். இவ்வாறு குறை திரிபுற்ற வினைகள் குறை வினைகள் எனப்படும். பின்வருவன தற்கால தமிழில் உள்ள குறை வினைகள் ஆகும்: *மாட்டு*, *உள்*, *இல்*, *வேண்டு*, *போது*, *தகு*, *முடி*, *தெரி*, *புரி*, *பிடி*, *கிடை*, *வலி*, *பசி*, *கூசு*, *கச*.

1. *மாட்டு* என்ற வினை துணை வினையாக இட-எண்- பால் ஒட்டுக்கள் பெற்று முற்று வினையாக மட்டும் வரும் (*மாட்டேன்*, *மாட்டாள்*, *மாட்டான்*).

3. *உள்* என்ற வினை சொல் வினையாகவும் துணை வினையாகவும் வரும். இது முற்று வடிவிலும் ஒரு எச்ச வடிவிலும் வரும். அதாவது *உண்டு* என்ற முற்று வடிவிலும் உள்ள என்ற எச்ச வடிவிலும் வரும்.

4. *இல், அல்* என்ற எதிர்மறை வினைகள் காலக் குறிப்பின்றி படர்க்கையிடப் பன்மை எண் மற்றும் அஃறிணைப்பால் என்பதற்குத் திரிபுற்று முறையே *இல்லை*, *அல்ல* என்ற வடிவங்களில் வரும். எதிர்மறைப் பெயரெச்ச வடிவங்களில் அதாவது *இல்லாமல்*, *அல்லாமல்* அல்லது *இன்றி*, *அன்றி* எனவும் வரும்.

5. *வேண்டு*, *போது*, *தரு* என்பன ஒரு முற்று வடிவில் எதிர்காலப் படர்க்கையிட ஒருமை என்பதற்குத் திரிபுற்று *வேண்டும்*, *போதும்*, *தகும்* என வரும். இவற்றிற்கு இணையான எதிர்மறை வடிவிலும், அதாவது *வேண்டாம்*, போதாது, *தகாது* என்றும் வரும். இவ்வினைகள் பின்வரும் வாக்கிய அமைப்பில் வரும்:

நான்காம் வேற்றுமைத்தொடர் + முதல் வேற்றுமைத்தொடர் + வினை

எ.கா.

குமாருக்குப் பணம் வேண்டும்

ராதைக்குத் தண்ணீர் வேண்டாம்

6. *முடி*, *தெரி*, *புரி*, *பிடி*, *கச* என்ற வினைகள் முற்றுப்பெறா வினைகள் வடிவிலும் முற்றுப்பெற்ற வினை வடிவிலும் மூன்று காலத்துக்கும் படர்க்கை இடத்துக்கும் ஒருமை பன்மை எண்ணுக்கும் அஃறிணை பாலுக்கும் திரிபுற்று வரும். *முடி* துணைவினையாகவும் பிறவினைகள் சொல் வினைகளாகவும் வரும். அவை பின்வரும் வாக்கிய ஒழுங்கமைப்பில் வரும்.

நான்காம் வேற்றுமைத் தொடர் + இரண்டாம் வேற்றுமைத்தொடர் + வினை

எ.கா.

ராதைக்குப் பாட்டு பிடிக்கும்

3.4 கணினி உருபனியல் ஆய்வு

உருபனியல் ஆய்வுகள் சேர்த்தல் (combinatorial) அல்லது உறவுப்படுத்தல் (relational) அணுகுமுறைகளை உள்ளடக்கியுள்ளன. சொல்லும் வரிசை முறையும் (Item and Arrangement (IA)) (Hockett 1958) ,சொல்லும் செயற்பாங்கும் (Item and Process (IP) (Hockett, 1958) போன்ற மாதிரிகள் ஒரு உருபனியல் கோர்வையை அடி (base) மற்றும் ஒன்று அல்லது ஒன்றுக்கு மேற்பட்ட ஒட்டுகளாகப் பிரிப்பதை உள்ளடக்கும். மாறாகச் சொல்லும் அடுக்கும் (Word and Paradigm (WP)) மாதிரி (Robbins, 1963) சேர்த்தல் மற்றும் உறவுப்படுத்தல் என்ற இரு அணுகுமுறைகளையும் உருப்படுத்தம் செய்யும். சொல் வலைப்பின்னல் (Word Network, Singh and Ford, 1984) போன்ற மாதிரிகள் உறவுப்படுத்தல் அணுகுமுறையைப் பண்பாக்கம் செய்யும். இது சொல்லை அலகாகக் கொள்கிறது; உருபனை அலகாகக் கொள்வதில்லை; உருபனியல் என்பது சொல் அதைப்போன்ற பிற சொற்களுடன் கொண்டுள்ள கலவைத் தன்மையான உறவாக விளக்கிக் கொள்ளப்படுகிறது.

சொல்லும் வரிசை முறையும் மாதிரியை அடிப்படையாகக் கொண்ட உருபனியல் ஆய்வில் எல்லா உருபன்களும் சில அலகுகள் என்று எடுத்துக்கொள்ளப்படும். மேலும் மொழியின் எல்லா உருபன்களையும் தேவையான பண்புக்கூறு குறிப்புடன் பட்டியலிடும். பகுதிகளும் ஒட்டுகளும் அவற்றின் தொடர்புபடுத்த இயலும் இணைகளைப் பற்றிய தகவலுடன் தனியாகச் சேகரிக்கப்படும்; சொல் வடிவுகளின் ஆய்வு பலவகைப்பட்ட கலவைத் தன்மை கொண்ட சொல் வடிவுகளில் உள்ள உருபனியல் வடிவமைப்புகளின் அடையாளத்தால் குறிப்பிடப்படும் அமைப்பொழுங்குப் பொருத்தம் பார்த்தல் (pattern matching) வழி செயல்படுத்தும்.

உருபனியல் ஆய்வில் சொல்லும் செயற்பாங்கும் உருபனியல் மாதிரியைப் பயன்படுத்தினால் இது செய்யும் கோட்பாடு அடிப்படையிலான அனுமானங்கள் அடிப்படையில் அருவத்தன்மையான உள்ளுறையும் உருபன் சூழல் அடிப்படையில் தீர்மானிக்கப்பட்ட உருபொலியனியல் நிலைகளில் வேறுபட்ட உருபுகள் வழியாக அதன் புற உருப்பத்தத்தைக் கண்டுகொள்கிறது. வேர்களும் ஒட்டுகளும் அவற்றின் அடிப்படை மாற்றுருபு வடிவுகளில் பட்டியலிடப் பட்டுள்ளன; மற்றும் வேறுபட்ட புற மாற்றுருபச் சேர்க்கைகளின் கட்டுப்பாடுகளைக் குறிப்பிடும் உருபனியல் விதிகளின் செயல்பாட்டின்வழி அணுகப்படுகிறது.

சொல் வடிவுகளை ஆய சொல்லடுக்கு மாதிரியைப் பயன்படுத்துவது உறவுபடுத்தல் அணுகுமுறையாகும். ஒவ்வொரு அடுக்கின் உருப்படுத்தங்களின் பட்டியல் போதுமானதல்ல; ஏனென்றால் ஒவ்வொரு அடுக்கின் செல்லுக்குள்ளும் செயல்படும் ஆக்கச் செயற்பாடுகள் தன்னிச்சையான நடத்தையை வெளிப்படுத்துகிறது. வேர்கள் மட்டும் சொல் பதிவுகளாக எடுத்துக்கொள்ளப்படுகிறது; வேறுபட்ட வேர்கள் மற்றும் பகுதிகள் மீதான எல்லாச் சாத்தியமான ஒட்டாக்கங்களும் சொல்லாக்க விதிகளின் அறிமுகம் வழி சாத்தியமாக்கப்படுகிறது. சொல் வடிவுகளின் ஆய்வு இச்சேர்ப்புத் தொடர்புகளின் மீது செயல்படும் விதிகளால் ஊக்கப்படுத்தப்படுகிறது.

சொல் மற்றும் வலைப்பின்னல் மாதிரி உருபனியல் ஆய்வுக்கு முழுப்பட்டியல் மாதிரியாக (Full Listing Model) கருதப்படுகிறது. உறவுப்படுத்தல் அணுகுமுறையின் முக்கியத்துவம் சொல் வடிவுகளின் ஆக்கத்திற்கும் புரிந்து கொள்ளலுக்கும் கூடுதல் உகந்ததாகத் தெரிகிறது. ஒட்டுகள் சொற்களுக்கு வெளியே எங்கும் இல்லாதிருப்பதால்

சொல்லாய்விக்கு இக்கோட்பாட்டைக் கையாள்வது இதை ஆக்கச் செயன்மைகளின் தொடர்ச்சிகளாகப் பார்ப்பதற்குப் பதிலாக வடிவுகளுக்கு இடையில் உள்ள வலைப்பின்னலின் ஒப்புமைப் படுத்தலை முற்கோள் செய்கிறது. வேர் அல்லது ஒட்டு என்ற கருத்துகளுக்குக் குறிப்பிட்ட உருபனியல் தகுதி தரப்படவில்லை. இருப்பினும் தொழில்நுட்ப நோக்கத்திற்கு வேண்டி இவ்வேறுபாடுகளைக் கருத்தில் கொள்வது அவசியமாகும்.

ஆம்பிள் (Ample) மற்றும் கிம்மோ (Kimmo) போன்ற உபாயங்கள் நன்றாக நிறுவப்பட்ட கணினியியல் வடிவவாதங்களாகும். இவை பல மொழிகளுக்கு உருபனியல் ஆய்விகளையும் உருவாக்கிகளையும் திட்டமிடப் பயனுள்ளதாக அமைகின்றன. ஒரு அர்த்தத்தில் ஆம்பிள் சொல்லும் வரிசை முறையும் மாதிரி அடிப்படையில் திட்டமிடப்பட்டிருந்தாலும் உள்ளுறையும் கொள்கையை மாற்றுருபுப் பொருத்தமாக (allomorph matching) விளக்கலாம். மொழித் தரவு மையம் மொழியின் சொல் வகுப்புகளின் ஒவ்வொன்றிற்கும் எல்லா மாற்றுருபுகளின் விரிவான பட்டியலைக் கொண்டிருக்கிறது. பகுதிகள் மற்றும் பின்னொட்டுகளின் விரிவான பட்டியலைக் கொண்டிருக்கிறது. பகுதிகள் மற்றும் பின்னொட்டுகளின் தவறான இணைப்புகளை ஒலியனியல், உருபொலியனியல் அல்லது உருபனியல் கட்டுப்பாடுகளை தேவையான இடத்தில் குறிப்பிடுவதன் மூலம் தடை செய்யலாம்.

கிம்மோ ஆக்கமுறை ஒலியனியல் அடிப்படையில் திட்டமிடப்பட்டுள்ளது. இது இருநிலை உருபனியல் (Two level Morphology) என்று அழைக்கப்படும் உள்ளுறையும் அடிப்படையை மாற்றுருபன்களைப் புற எழுத்து வடிவுகளாக மாற்றும் இருதிசை பகுப்பாய்வியாக (bi-directional processor) நிறைவேற்றப்படுகிறது. ஒரு ஒலியனியல் விதி சொல் வடிவுகள் வரும் சூழல்களின் அடிப்படையில் சொல்லின் உள்ளுறை வடிவிற்கும் அதன் புற மெய்படுத்தத்திற்கும் இடையிலுள்ள நேரடிப் பொருத்தத்தை உருவாக்குகிறது. விதிகள் பயன்பாட்டின் நிகழ்வெண்ணைக் குறிப்பிடுவதுடன் அவை வரும் சூழலின் அடிப்படையில் மாற்று விதிகளின் இயல்பைத் தொடர்புபடுத்துகிறது. அதன் செயல்பாடு முற்றுநிலைத் தானியங்கியை (Finit State Automaton) சார்ந்திருக்கிறது. இது உள்ளீடு செய்யப்படும் சொல்லைத் தரப்பட்ட கோர்வையின் தொடர்ச்சியான வருடி உணர்தல் (scanning) மூலம் பகுப்பாய்கிறது.

கோட்பாடு அடிப்படையில் நிறுவப்பெற்ற கொள்கைகளின் குழுமத்தைக் கணினிச் சட்டகத்தில் பொருந்துமாறு அப்படியே மாற்ற இயலாது. உருபனியல் வகுப்புப் பண்புக்கூறுகளை அல்லது செயற்பாங்குகளை ஒரு குறிப்பிட்ட வடிவ வாதமாக நடைமுறைப்படுத்தப்படுவது, நல்லுருவாக்கப்பட்ட நுட்பமான செயல்பாட்டு மாதிரியாக வரவேண்டி கோட்பாட்டின் சில தொழில்நுட்ப மாற்றியமைத்தலை வேண்டும். உருபனியல் ஆய்வின் மாதிரிகளின் வேறுபட்ட நடைமுறைப்படுத்தங்கள் கொள்கையும் வடிவாதமும் ஒன்றுக்கொன்று பொருத்தமாகச் செல்கிறதா என்பதைப் புரிந்துகொள்ள உதவும்.

3.4.1 கலப்பு மாதிரி

பெர்லில் (PERL) இல் உருவாக்கப்பட்ட மற்றும் செயல்படுத்தப்பட்ட கலப்பு மாதிரி (Hybrid Model) உள்ளுருவாக்கப்பட்ட சொற்களஞ்சியத்தின் சட்டகத்தையும் மொழியின் வேறுபட்ட ஒவ்வொரு சொல் வகுப்பிற்கும் சொல்லும் வரிசை முறையும் மற்றும் சொல்லும் செயற்பாங்கும் என்பனவற்றின் ஒருகுழும விதிகளையும் கொண்டிருக்கிறது. தரப்பட்ட உருபொலியனியல் விதிகளின் அடிப்படையில் செயல்படும் இணையான துணை நடைமுறைகளை (corresponding subroutines) வேண்டுவதன் மூலம் இதன் ஆய்வு ஆய்வு நிரூபிக்கப்படுகிறது. இது ஒவ்வொரு வகைப்பாட்டின் சொல் வடிவுகளைக் கோப்புகளால் கட்டளையிட்டபடி செயல்படுகிறது. வழியமைப்பு மொழியில் மொழித் தகவலின் மரபுரிமை குறியாக்கம் இருக்கிறது. திறந்த இறுதியுற்ற உருபொலியனியல் மற்றும் மாற்றுருப விதிகளின் (open ended morphophonemic and allomorphy rules) வழி மொழித் தகவலின் பகுப்பாய்வால் இது நடைபெறுகிறது (Vaishnavi Ramaswamy, 2003: 101-110).

இம்மாதிரி உருபனியலின் சொல்லும் வரிசை முறையும் மற்றும் சொல்லும் செயற்பாங்கும் மாதிரிகளின் கலவையில் அதன் கோட்பாடு அடிப்படையை அமைத்துக்கொள்கிறது. செயற்பாட்டைச் சொற்பகுதியின் (word stem) உருபனியல் இயல்பைப் பொறுத்து ஒரு மாற்றத்தை (transformation) விவரணை செய்வதாக விளக்க இயலும். திரிபு நீக்கத்தின் (Inflectional stripping) ஒவ்வொரு நிலையிலும்

குறிப்பிடப்பட்டுள்ள கட்டுப்பாடுகளின் முழுத் தொடர்ச்சிகளின் வழி கடந்த பின்னர் ஒரு சொல் பகுப்பாய்வு செய்யப்படுகிறது.

ஒரு பகுப்பாய்வுத் தொகுதி (analysis modeule) வலமிருந்து இடமாகக் கோர்வையை வருடி அறிதலைச் (scanning) செய்கிறது; ஒரு நேரத்தில் ஒரு பின்னொட்டை வருடி அறிந்து பிரித்தெடுக்கிறது; மற்றும் ஒவ்வொரு நேர்விலும் ஒலியனியல் மற்றும் உருபொலியனியல் உதவியுடன் மீதிச்சொல்லை மீட்டுருவாக்கம் செய்கிறது. கோர்வை தீர்வது வரை இது நடைபெறும்; வெளியீடு கண்டுபிடிக்கப்பட்ட உருபன்களின் வடிவில் இருக்கும்; பின்னர் அவை உறுதி செய்யப்படும்.

3.4.1.1 கோட்பாடும் செயல்படுத்தமும்

இம்மாதிரி உருபனியலின் சொல்லும் வரிசையும் அணுகுமுறையை நோக்கி அமைக்கப்பட்டுள்ளது. புணர்ச்சி (sandhi) போன்ற இணைப்பு சாரா நடப்புக்கான (Non-concatinative phenomenon) தீர்வுகளுக்குச் சொல்லும் செயற்பாங்கும் அணுகுமுறையைக் கையாள்கிறது. ஒவ்வொரு ஒலியனியல் அல்லது உருபொலியனியல் நடப்பும் மாற்றுவிதிகளில் (transformational rules) உட்படுத்தப்படுகிறது. உருபனியல் அடுக்கின் ஒவ்வொரு அலகையும் உள்ளடக்கிய உருபொலியனியல் மாற்றங்கள் இடப்பெயர்ச்சி அல்லது மாற்ற விதிகளின் ஒரு குழுமத்தின் மூலம் எடுத்துக்காட்டப்படுகின்றது.

3.4.1.2 வடிவ வாதமும் கட்டும்

பெர்லில் (PERL) நடைமுறைப்படுத்தப்படும் கலவை மாதிரி ஒரு கணினிசார் மாதிரியாகும். இது சுதந்திரமாகச் செயல்படும் ஐந்து தனிப்பகுதிகளாகச் (separate blocks) செயல்படும் சிறு தொகுதிகளை வேண்டும் முதன்மைக் கட்டுப்பாட்டு வழியமைப்புகள் /நிரல்கள் (main control programmes) வழியாகச் செயன்மை புரியும் ஒருங்கிணைந்த மாதிரியாகச் செயல்பட இயலும். கட்டுப்படுத்தும் தொகுதி (controlling modules) கட்டுப்படுத்தும் தகவலின் பகுதியாகப் பின்வரும் ஆறு பகுதிகளை வேண்டும்:

1. பெயர்த் தமிழ் - Noun Tamil
2. மாற்றுப்பெயர்த் தமிழ் - Pronoun Tamil

3. எண் தமிழ் - Numeral Tamil
4. வினைத் தமிழ் - Verb Tamil
5. பெயரடைத் தமிழ் - Adjective Tamil
6. சொற்களஞ்சியம் - Lexicon
7. வெளியேறும் வழி - Exit

1-இலிருந்து 5 வரையுள்ள பாகங்கள் தமிழில் ஒவ்வொரு சொல் வகுப்பிற்கும் ஆய்வு உட்கூறை (analysis component) உட்படுத்தும். ஒவ்வொரு பாகமும் பொருத்தமான சொல் வகுப்பின் ஆய்வுக்கு விதிகளை உட்படுத்தும். ஆய்வு உள்ளீட்டுக் கோவையின் வலது முனையிலிருந்து தொடங்கி ஒவ்வொரு மட்டத்திலும் இறுதிப் பின்னொட்டை அல்லது திரிபைப் பிரித்தெடுப்பதால் தொடர்ந்து நடைபெறும்.

கட்டளைக் கோப்பு (control file) வழியமைப்பில்/நிரலில் (programme) உள்ள அகக் கட்டளைகளை வேண்டும். உள்ளீடாக எடுக்கப்பட்ட எந்தச் சொல் வடிவமும் ஒவ்வொரு சொல் வகுப்பையும் (அதாவது பெயர், மாற்றுப்பெயர், எண், வினை மற்றும் பெயரடை) கையாள உருவாக்கப்பட்ட ஐந்து ஆய்வுத் தொகுதிகளால் தொடர்ச்சியான பரிசோதனைகளுக்கு உட்படுத்தப்படும். வெற்றிபெறும் பகுப்பாய்வு வெளியீடாக விளைவிக்கப்படும்.

லெக்சிக்கன் (LEXICON) என்ற சொற்களஞ்சியக் கோப்பு (lexicon file) சொல் வடிவுகளைத் தொடரும் எல்லாத் திரிபுகளின் மற்றும் பின்னொட்டுகளின் விவான பட்டியலைக் கொண்டிருக்கும் உட்கூறு ஆகும். திரிபுகள் மாற்ற உட்கூறில் (transformational component) மாற்றொலிக் கட்டுப்பாடுகளாகக் (allophonic conditionings) கருதப்படுவதால் பின்னொட்டுக்கள் சொல் பதிவுகளாகப் பதிவுசெய்யப்பட வேண்டும். எடுத்துகாட்டாக உடைமை வேற்றுமையைக் குறிக்கும் *உடைய* என்ற கோர்வை ஒரு குழும மாற்றங்களைச் செயல்படுத்தும் கண்ணிக்கு முக்கியச் சொல்லாகும் (Key word). இம்மாதிரியில் தொகுதியில் (module) உள்ள எல்லாத் தகவல்களும் வழியமைப்புக்குள் கையாளப்படுகிறது.

உருபனின் வகைப்பாடு, அதன் ஒலியனியல் மற்றும் உருபனியல் பண்புகள் மற்றும் அதனைப் பிணைக்கும் கட்டுப்பாடு பற்றிய தகவல்கள் எல்லாம் அடையாளப்படுத்தப்பட்ட மாறிகளாகத் (tagged variables) தரப்படுகிறது. உருபொலியனியல் மற்றும் உருபன் அமைப்பொழுங்கு (morphotactic) கட்டுப்பாடுகள் கண்ணிகளின் (looping) நிரல் வழியாகவும் கண்ணிகளுக்குள் விதிகளின் நிரல் வழியாகவும் புகுத்தப்படலாம். ஒலியன் அமைப்பொழுங்கு (phonotatics) மற்றும் ஒலியனியல் கட்டுப்பாடுகள் கண்ணிக்குள் ஒரு குறிப்பிட்ட உருபனுக்குச் சிறப்பீடுகளாகப் புகுத்தப்படும்

கோர்வைச் சூழ்நிலைக் கட்டுப்பாடுகள் (String Environment Constraints = SECs) ஒட்டின் ஒவ்வொரு மாற்றுருபுக்கும் பிரயோகிக்கப்படுகிறது. குறிப்பாக முக்கியமான உருபன் எல்லை குறிப்புகளில் (leading morpheme boundary markers) பிரயோகிக்கப்படும். உருபன் சூழ்நிலைக் கட்டுப்பாடுகள் (Morpheme Enveronment Constraints = MECs) தற்போதைய உருபனின் மாற்றுருபன் மீது உருபன் சூழ்நிலைக் கட்டுப்பாடுகளைப் புகுத்தும்.

3.4.1.3 கலவை மாதிரிக்கான சட்டகம்

தற்போதைய மாதிரி போன்ற வடிவ வாதத்தில் தமிழின் உருபனியல் வகுப்புகளின் கலவைத் தன்மையான ஒட்டுநிலை உருபனியலை (complex agglutinative morphology) அறிந்து கொள்வது பொருத்தமான சேர்க்கைக் கட்டுப்பாடுகளுடன் (relevent combinatirial consrraints) பகுதிகள் (stems) மற்றும் பின்னொட்டுகள் என்பனவற்றின் மாற்றுருபுகளின் விரிவானப் பட்டியலை வேண்டும். இம்மாதிரியான சொல்லும் வரிசை முறையும் மாதிரியிலிருந்து (IA model) அதன் அடிப்படைக் கொள்கைகளைப் பெறுவதற்கு இவ்வாறு அமைகின்றது. மேக்ஸ்வெல்லினுடைய (Maxwell’s) பார்வையின்படி கோட்பாடு அடிப்படையில் சொல்லும் வரிசை முறையும் மாதிரியும் சொல்லும் செயற்பாங்கும் மாதிரியும் (IP model) அகன்ற வேறுபாடுகளை உருப்படுத்தத்தினாலும் பயன்பாட்டில் அக்கொள்கைகள் இணைகின்றன.

3.4.1.4 மாற்றுருப நிலை

பகுதி மற்றும் பின்னொட்டு மாற்றுருப நிலையில் (sterm and suffix allomorphy) சிக்கனத்தைப் பற்றிய அக்கறைக்கு முன் சிந்தனை தரப்பட வேண்டும். மாற்றுருப நிலையின் எளிமையாக்கம் வடிவ வாத்தில் சிக்கனத்தை எடுத்துக்காட்டுகிறதா, அப்படியென்றால் சொல்லின் எப்பாகம் உருபனியல் அடிப்படையில் எளிமையாக்கம் செய்யப்பட வேண்டும் என்பன பற்றிய தீர்மானம் எடுக்கப்பட வேண்டும். உள்ளுணர்வு அடிப்படையில் மாற்றுருப நிலையின் சுமையைப் பகுதியிலிருந்து பின்னுருபுக்கோ அதன் மறுதலைக்கோ மாறுதல்செய்வது சரியானது என்று தீர்மானிக்கக்கூடியதாக இருந்தால் நடைமுறைக்குப் பொருத்தமாகச் சொல் வகுப்புகளின் அடுக்குமுறை அமைப்புகள் (paradigmatic struture) மாற்றப்பட வேண்டும்.

மாற்றுருப நிலையின் சிக்கனத்தை விதிகள் மற்றும் அல்லது கட்டுப்பாடுகளின் சிக்கனத்தின் மீது சலுகை செய்து கொள்ளலாம். சொல்லும் வரிசையும் சட்டகத்தில் உருபனியலைத் திரும்பப்பெறும் நல்வழி மொழியின் உருபனியல் வடிவுகளின் எல்லா மாற்றுருப மாறிகளையும் (allomorphic variants) பட்டியலிடுவதுதான். நடைமுறை நோக்கத்திற்காக அடுக்குகளின் கோட்பாட்டுக் குழுமங்களை நேரடியாக எடுத்துக்கொள்ள இயலும்.

3.4.1.4.1 பெயர்கள்

பெயர்ச் சொற்களுக்குக் குறைந்து 23 அடுக்கு வகைகளை நிறுவ வேண்டும். அவை கீழே பட்டியலிடப்பட்டுள்ளன.

வ.எண்	**பகுதி/பின்னொட்டு[1] (அகராதிப் பதிவு)**	**பகுதி/ பின்னொட்டு[2]**	**பகுதி/ பின்னொட்டு[3]**
1.	ஆண்	ஆண்] Ø	ஆண்]கள்
2.	ஆறு	ஆற்]ற்	ஆறு]கள்
3.	எலி	எலி]ய்	எலி]கள்
4.	இலை	இலை]ய்	இலை]கள்
5.	இதழ்	இதழ்] Ø	இதழ்]கள்

6.	ஈ	ஈ]ய்	ஈக்]கள்
7.	கண்	கண்]ண்	கண்]கள்
8.	கால்	கால்] Ø	கால்]கள்
9.	காடு	காட்]ட்	காடு]கள்
10.	மனிதன்	மனிதன்] Ø	மனிதர்]கள்
11.	மரம்	மர]த்த	மரங்]கள்
12.	முள்	முள்]ள்	முட்]கள்
13.	மான்	மான்] Ø	மான்]கள்
14.	நரம்பு	நரம்ப்] Ø	நரம்பு]கள்
15.	நாய்	நாய்] Ø	நாய்]கள்
16.	பொன்	பொன்]ன்	பொன்] Ø
17.	பொய்	பொய்]ய்	பொய்]கள்
18.	புல்	புல்]ல்	புற்]கள்
19.	பொருள்	பொருள்] Ø	பொருட்]கள்
20.	பூ	பூ]வ்	பூ]க்கள்
21.	வண்டு	வண்ட்] Ø	வண்டு]கள்
22.	தேர்	தேர்] Ø	தேர்]கள்
23.	தோள்	தோள்] Ø	தோள்]கள்

விளைகின்ற மாற்றுருபுகளின் மொத்த எண்ணிக்கை பின்வருமாறு அமையும்: 23 அடிப்படைப் பகுதிகள் +10 பகுதி மாற்றுருக்கள் (stem variants) +9 திரிபுக் குறிகள் (oblique markers) (x 8 வேற்றுமைப் பின்னொட்டுகள்) +8 அடிப்படை வேற்றுமைப் பின்னொட்டுகள் +1 எண் பின்னொட்டு மாற்றுரு = 115 பெயர்களின் மாற்றுருபுகள்.

3.4.1.4.2 மாற்றுப்பெயர்கள்

நாம் மற்றுப்பெயருக்குக் குறைந்தது 15 அடுக்கு வகைகளை நிறுவ வேண்டும். அவை கீழே பட்டியலிடப்பட்டுள்ளன.

வ.எண்	**பகுதி1/ பின்னொட்டு**	**பகுதி2/ பின்னொட்டு**	**பகுதி3/ பினொட்டு**
1.	அவள்	அவள்] Ø	அவள்] Ø
2.	அவன்	அவன்] Ø	அவன்] Ø
3.	அவர்	அவர்] Ø	அவர்
4.	அது	அத்] Ø	அத்]த்
5.	அவை	அவற்] Ø	அவற்]ற்
6.	நான்	என்] Ø	என்]ன்
7.	நாங்கள்	எங்கள்] Ø	எங்கள்] Ø
8.	நீ	உம்] Ø	உம்]ம்
9.	நீங்கள்	உங்கள்] Ø	உங்கள்] Ø
10.	நீர்	உம்] Ø	உம்]ம்
11.	யாம்	எம்] Ø	எம்]ம்
12.	தாங்கள்	தங்கள்] Ø	தங்கள்] Ø
13.	தாம்	தம்] Ø	தம்]ம்
14.	தான்	தன்] Ø	தன்]ன்
15.	நாம்	நம்] Ø	நம்]ம்

விளைகின்ற மாற்றுருபுகளின் மொத்த எண்ணிக்கை பின்வருமாறு அமையும்: 15 அடிப்படைப் பகுதிகள் + 11 பகுதி மாற்றுருக்கள் +8 அடிப்படை வேற்றுமைப் பின்னொட்டுகள் +3 திரிபுக் குறிகள் [x 8 வேற்றுமைப் பின்னொட்டுகள்] = 58 மாற்றுப் பெயர்களின் மாற்றுருபுகள்.

3.4.1.4.3. எண்கள்

நாம் எண்களுக்குக் குறைந்தது 14 அடுக்கு வகைகளை நிறுவ வேண்டும். அவை கீழே பட்டியலிடப்பட்டுள்ளன:

வ.எண்	பகுதி1/ பின்னொட்டு1	பகுதி2/ பின்னொட்டு2	பகுதி3/ பின்னொட்டு3	பகுதி4/ பின்னொட்டு4
1.	ஒன்று	ஒன்ற்] Ø	ஒரு]வ்	ஒன்ற்]ஆம்
2.	இரண்டு	இரண்ட்] Ø	இரு]வ்	இரண்]ஆம்
3.	மூன்று	மூன்ற்] Ø	மூ]வ்	மூன்ற்]ஆம்
4.	நான்கு	நான்க்] Ø	நால்]வ்	நான்க்]ஆம்
5.	ஐந்து	ஐந்த்] Ø	ஐ]வ்	ஐந்த்]ஆம்
6.	ஆறு	ஆற்] Ø	ஆறு]வ்	ஆற்]ஆம்
7.	ஏழு	ஏழ்] Ø	ஏழு]வ்	ஏழ்]ஆம்
8.	எட்டு	எட்ட்] Ø	எண்]ம்	எட்ட்]ஆம்
9.	ஒன்பது	ஒன்பத்] Ø	ஒன்பது] Ø	ஒன்பத்]ஆம்
10.	பத்து	பத்த்] Ø	பத்து] Ø	பத்த்]ஆம்
11.	நூறு	நூற்]ற்	நூறு] Ø	நூற்]ஆம்
12.	ஆயிரம்	ஆயிர]த்த்	ஆயிரம்] Ø	ஆயிர]த்து
13.	லட்சம்	லட்ச]த்த்	லட்சம்] Ø	லட்ச]த்து
14.	கோடி	கோடி]ய்	கோடி] Ø	கோடி] Ø

விளைகின்ற மாற்றுருபுகளின் மொத்த எண்ணிக்கை பின்வருமாறு அமையும்: 14 அடிப்படை பகுதிகள் + 13 பகுதி மாற்றுருக்கள் +3 முறைமைக்குறி பின்னொட்டுகள் +2 பால் – எண் – இடக்குறி மாற்றுருக்கள் (x 1 பால்- எண்- இடப் பின்னொட்டு) + 3 பால் – எண் – இடப் பின்னொட்டுகள் + 8 அடிப்படை வேற்றுமைப் பின்னொட்டுகள் + 3 திரிபுக்குறிகள் (x 8 வேற்றுமைப் பின்னொட்டுகள்) = 67 எண் மாற்றுருபுகள்.

3.4.1.4.4 வினைகள்

நாம் வினைச்சொல்லுக்குக் குறைந்தது 24 அடுக்குமுறைகளை நிறுவ வேண்டும். அவை கீழே பட்டியலிடப்பட்டுள்ளன.

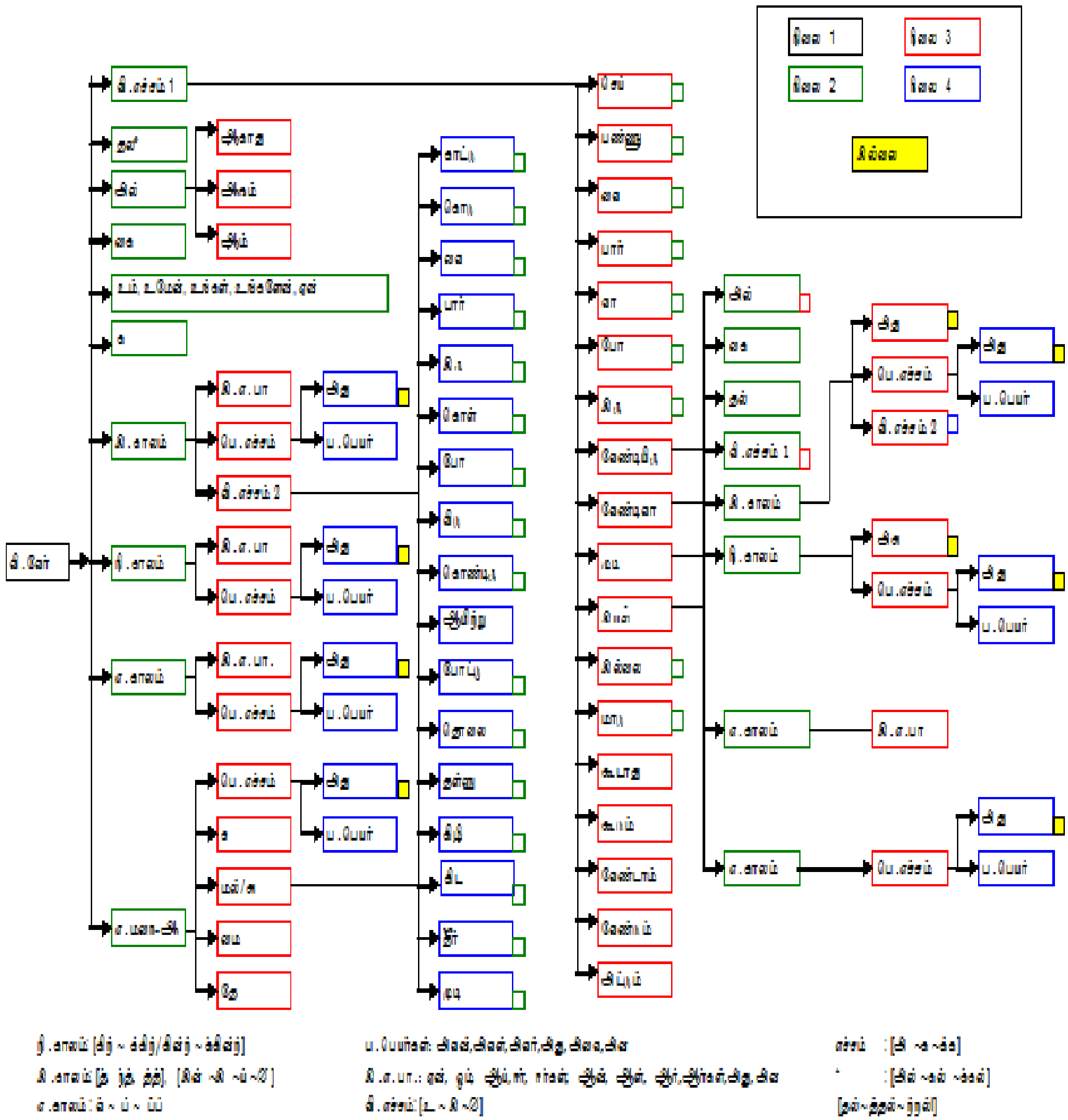
வ.எண்	அடி1/ பின்னொட்டு	அடி2/ பின்னொட்டு	அடி3/ பின்னொட்டு (நிகழ்காலம்)	அடி4/ பின்னொட்டு (எதிர்காலம்)	அடி5/ பின்னொட்டு (முடிவுறா எச்சம்)
1.	ஆகு	ஆ]ன்	ஆ]கிற்	ஆ]வ்	ஆக்]அ
2.	ஆள்	ஆண்]ட்	ஆள்]கிற்	ஆள்]வ்	ஆள்]அ
3.	அழு	அழு]த்	அழு]கிற்	அழு]வ்	அழ்]அ
4.	சா	செ]த்த	சாகி]ற்	சா]வ்	அசா]க
5.	செல்	சென்]ற்	செல்]கிற்	செல்]வ்	செல்]ல
6.	செய்	செய்]த்	செய்]கிற்	செய்]வ்	செய்]ய
7.	சொல்	சொன்]ன்	சொல்]கிற்	சொல்]வ்	சொல்]ல
8.	என்	என்]ற்	என்]கிற்	என்]ப்	என்]அ
9.	இகழ்	இகழ்]ந்த்	இகழ்]கிற்	இகழ்]வ்	இகழ்]அ
10.	இரு	இரு]ந்த்	இரு]க்கிற்	இரு]ப்ப்	இரு]க்க
11.	காண்	கண்]ட்	காண்]கிற்	காண்]ப்	காண்]அ
12.	கேள்	கேட்]ட்	கேட்]கிற்	கேட்]ப்	கேட்]க
13.	கொள்	கொண்]ட்	கொள்]கிற்	கொள்]வ்	கொள்}]ள
14.	நில்	நின்]ற்	நிற்]கிற்	நிற்]ப்	நிற்]க
15.	நோ	நொ]ந்த்	நோ]கிற்	நோ]வ்	நோ]க
16.	படி	படி]த்த்	படி]க்கிற்	படி]ப்ப்	படி]க்க
17.	பெறு	பெற்]ற்	பெறு]கிற்	பெறு]வ்	பெற்]அ
18.	போ	போ]ன்	போ]கிற்	போ]வ்	போ]க
19.	போடு	போட்]ட்	போடு]கிற்	போடு]வ்	போட்]அ
20.	பூண்	பூ]ட்	பூணு]கிற்	பூணு]வ்	பூண்]அ
21.	உண்	உண்]ட்	உண்]கிற்	உண்]ப்	உண்]ண
22.	வா	வ]ந்த்	வரு]கிற்	வரு]வ்	வர்]அ

23.	விழு	விழு]ந்த்	விழு]கிற்	விழு]வ்	விழ்]அ
24.	தூங்கு	தூங்க்]இன்	தூங்கு]கிற்	தூங்கு]வ்	தூங்க்]அ

விளைகின்ற மாற்றுருபுகளின் மொத்த எண்ணிக்கைப் பின்வருமாறு அமையும்: 24 அடிப்படை வேர்கள் + 33 பகுதி மாற்றுருக்கள் +12 காலக்குறிகள் + 7 எச்சக்குறிகள் + 63 முற்று வினையெச்ச அடிப்படையிலான பின்னொட்டு மாற்றுருக்கள் + 31 முற்று / கால அடிப்படையிலான பின்னொட்டு மாற்றுருக்கள் = 170 வினை மாற்றுருபுகள்.

வினைக் கூட்டமைப்பின் ஒழுக்கு

முனைவர் ச.இராசேந்திரன், தமிழ்ப்பல்கலைக்கழகம், தஞ்சாவூர்



3.4.1.4.5 பெயரடைகள்

பெயரடைகளுக்கு நாம் குறைந்தது இரண்டு அடுக்குமுறை வகைகளை நிறுவ வேண்டும். அவை கீழே பட்டியலிடப்பட்டுள்ளன.

வ.எண்	பகுதி$_1$/ பின்னொட்டு$_1$	பகுதி$_2$/ பின்னொட்டு$_2$	பகுதி$_3$/ பின்னொட்டு$_3$
1.	நல்ல	நல்ல] Ø	நல்ல]வ்
2.	புது	பு]இய	புதிய]வ்

விளைகின்ற மாற்றுருபுகளின் எண்ணிக்கைப் பின்வருமாறு அமையும்: 2 அடிப்படைப் பகுதிகள் + 1 பகுதி மாற்றுரு +1 சாரியைப் பின்னொட்டு + 1 பால் – எண் – இடம் பின்னொட்டு +1 பால் – எண் – இடக்குறி (x 3 பால் – எண் – இடப் பின்னொட்டுகள், = 8 பெயரடை மாற்றுருபுகள்.

3.4.1.5 கலவை மாதிரியை நடைமுறைப்படுத்துதல்

ஐந்து வேறுபட்ட சொல் வகுப்புகளின் பகுப்பாடுக்குத் தேவையான துணை வழிகள் கலவை மாதிரியில் இருக்கின்றன. இம்மாதிரி கூடுதலாகச் சொல்லும் வரிசை முறையும் உருபனியல் மாதிரியில் செயல்படுவதால் பகுப்பாய்வின் ஒவ்வொரு நிலையினும் நேரிடும் புணர்ச்சி மாற்றத்தை வெளிப்படுத்தும் சொல் பகுதிகளின் அல்லது அடிகளின் முழு குழுமத்தின் ஒரு தெரிவடைவையும் (inventory) ஒலியனியல் அல்லது உருபொலியனியல் மாற்றங்களை விளக்கும் பொருத்தமான விதிகளையும், முன்வரும் பின்னொட்டு பிரித்தெடுக்கப்பட்டபின் மீதி இருக்கும் சொல் பகுதியின் மீட்டுருவாக்கத்தையும் தரவுமையம் (database) உட்படுத்தும்.

தற்போதைய நடைமுறைபடுத்தல் உருபனியல் பகுப்பாய்வுக்குச் சொல்லும் வரிசை முறையும் மாதிரியைப் பெரும்பாலும் பின்பற்றுகிறது. தரவு மையம் எல்லாப் பகுதிகளின் மாற்றுருப வேறுபாடுகளுடனும், மிகை ஆக்கத்தைத் (overgeneration) தடைசெய்ய சேர்க்கை விதிகளாக (combianatorial rules) அல்லது சூழல் காட்டுப்பாடுகளாகச் செயல்படும் கோர்வைச் சூழல் கட்டுப்பாடுகளுடனும் (string environment constraints)

உருபன் சூழல் கட்டுப்பாடுகளுடனும் (morpheme environment constraints) உருபன் சேர்ந்து வருகைக் கட்டுப்பாடுகளுடனும் (morpheme co-occurrence Constraints) இருக்கின்றது.

உளமொழியியல் வழியில் இம்மாதிரியின் உருபனியலானது கோர்வை வருடல் மாதிரிகளின் (String-Scan Models) எந்தச் சட்டகத்துடனும் பொருந்தவில்லை. இதை வகைப்பாடு-பிளவு மாதிரிகளுடன் (Category-Split Models) குழுமம் செய்யலாம். குறிப்பாக டாஃப்ட் மற்றும் ஃபாஸ்டர் மாதிரியுடன் (Taft and Foster Model) பொருத்தமுறும். தரப்பட்ட கோர்வைத் தொடர்ச்சியின் வருடல் இடமிருந்து வலமாகத் தொடரும். பின்னொட்டுகள் ஒத்தறி அடிப்படையில் பகுதிகளைவிடக் கூடுதலாக மாற்றொலியனியல் அல்லது மாற்றுருபன் மாற்றங்களுக்கு உள்ளாகும். எனவே இடமிருந்து வலமான பகுப்பாய்வு மிக எளிதாக மயக்கத்தைத் தீர்க்கும்.

இதன் விளைவாக மொழியில் உள்ள சொல் வகுப்புகளின் எண்ணிக்கையின் அடிப்படையில் இம்மாதிரியில் உள்ள ஆறு நடைமுறை ஒழுங்குகள் மொழிக்கென்று தனிப்பட்டவை (language specific) ஆகும். மொழித் தரவுக் குறியனாக்கத்திற்குத் (language data encoding) தமிழ் ஐந்து கோப்புகளையோ நடைமுறை ஒழுங்குகளையோ உட்படுத்துகிறது: பெயர்ச் சொற்களுக்கு, மாற்றுப் பெயர்களுக்கு, எண்களுக்கு, வினைகளுக்கு மற்றும் பெயரடைகளுக்கு என ஒவ்வொன்றுக்கும் ஒன்று. இதுதவிர மொழியின் எல்லாப் பின்னொட்டுகளின் ஒரு பட்டியலை வேண்டும் சொற்களஞ்சியப் பகுதி (Lexical Module) இருக்கிறது. பகுத்தாயப்பட வேண்டிய ஒவ்வொரு தரவுக்கோப்பும் திரிபு நீக்கலின் (infilectional stripping) குறிப்பிட்ட நிரலுக்கு இணங்க முன்னேறும் துணை நடைமுறை ஒழுங்காக (subroutine) வழியமைப்பு (programmed) செய்யப்பட்டுள்ளது.

வினையடைகள், பின்னுருபுகள், இடைச்சொற்கள், துணை வினைகள், குறைச் சொற்கள் போன்ற திரிபு நீட்சிகளின் வெளி மட்டங்கள் பொதுவான சொற்களஞ்சியத்தில் குறிப்பிடப்படுகிறது. வகுப்புச் சிறப்பீடுடைய (class-specific) எண், வேற்றுமை மற்றும் கால வடிவுகள் போன்ற திரிபுகள் பொருத்தமான வகைப்பாட்டுச் சிறப்பீட்டுத் துணை நடைமுறை ஒழுங்குகளில் குறிப்பிட்ட கண்ணியில் (specific loops) கூறப்படுகின்றன.

ஒவ்வொரு திரிபும்/பின்னொட்டும் சொற்களஞ்சியத்தில் பார்க்கப்பட்டு நீக்கப்படவும் அடையாளப்படுத்தப்படவும் தற்காலிக வரிசையில் சேகரிக்கப் படவும் செய்யப்படுகிறது. மீதியிருக்கும் பகுதி மீட்டுருவாக்கம் செய்யப்பட்டுத் திரிபு நீக்கலின் அடுத்த மட்டதிற்கு அனுப்பப்படுகிறது. உள்ளீடு செய்யப்பட்ட கோர்வையின் எல்லாச் சாத்தியமான திரிபுகளும் பின்னொட்டுகளும் தீர்க்கப்படுவது வரை இது நடைபெறும். இறுதியாக மீட்டுருவக்கப்பட்ட சொல் பகுதியுடன் கண்டறியப்பட்ட எல்லாத் திரிபுகளும் பின்னொட்டுகளும் இறுதி விளைவாக வெளிக்காட்டப்படும்.

எல்லா உருபன்களும் அப்படியே கூறப்படுவதால் மொழிக்கு ஒரு ஒலியனியல் தெரிவடைவு தேவையில்லை. உருபன் கோர்வைகள் அவை வரும் சம்பந்தப்பட்ட கண்ணிகளில் அடையாளப்படுத்தப்படுகின்றன.

திரிபு நீக்கலுக்கும் மீதி பகுதியின் மீட்டுருவாக்கத்திற்கும் உரிய ஒலியனியல் மற்றும் உருபொலியனியல் மாற்றங்கள் ஒவ்வொரு மட்டத்திலும் விதிகளாக உட்படுத்தப்படுகின்றன. கோர்வைக்கும் உருபன் வகுப்புகளுக்கும் கூடுதலான உட்கூறை உட்படுத்தும் ஆம்பிள் அல்லது கிம்மோ போல் அல்லாமல் இதில் ஒலியன் வரன்முறையையோ (phonotactics) உருபன் அமைப்பொழுங்கையோ (morphotactics) கையாள எதிர்மறை வகுப்பை குறியனாக்கம் செய்யும் விரிவான வழி இல்லை. இதற்குப் பதிலாகச் சூழ்நிலை மற்றும் சேர்ந்து வருகை கட்டுப்பாடுகள் ஒவ்வொரு கண்ணியிலும் குறிப்பிட்ட உருபன் கோர்வையில் நேரடியாக வைக்கப்பட்டுள்ளது. இது சொல் வகுப்பின் ஒவ்வொரு அடுக்குப் பதிவுக்கும் (pradigmatic entry) உரிய தேவையான மாற்றங்களை நடைமுறைப்படுத்துகிறது.

3.5 தமிழ்ப் புணர்ச்சி விதிகள்

தமிழ்ப் பனுவல்களின் பிழை திருத்தத்திற்குப் புணர்ச்சி விதிகளை அறிதல் முக்கியமானதாகும். பின்வருவன முக்கியமான புணர்ச்சி விதிகளாகும்.

1. *அந்த*, *இந்த* என்ற சுட்டுப் பெயரடைகளின் பின்பும் *எந்த* என்ற வினாப் பெயரடையின் பின்பும் ஒற்று மிகும்.

   அந்த + சாலை = அந்தச்சாலை

இந்த+ தலை = இந்தத் தலை

எந்த + பாய் = எந்தப் பாய்

2. *போக*, *வர*, *படிக்க*, போன்ற செய்ய-வினையெச்சங்களின் பின் ஒற்று மிகும்.

போக + கண்டேன் = போகக் கண்டேன்

வர + சொன்னான் = வரச் சொன்னான்

படிக்க + துணிந்தான் = படிக்கத்துணந்தான்

3. *சரிவர*, *கூட*, *மிக* போன்ற வினையடைகளின் பின்னும் *விட*, *போல* போன்ற இடைச் சொற்களின் பின்னும் ஒற்று மிகும்.

சரிவர + சொல் = சரிவரச் சொல்

கூட + போனான் = கூடப் போனான்

மிக + கவலைப்பட்டான் = மிகக் கவலைப்பட்டான்

அதைவிட + கூர்மையான = அதைவிடக் கூர்மையான

அவனைப் போல + பாய்ந்தான் = அவனைப் போல பாய்ந்தான்

4. ஆக என்ற வினையடை விகுதியின் பின் ஒற்று மிகும்

அழகாக + பேசு = அழகாகப் பேசு

5. *நல்ல*, *இன்ன*, *இன்றைய போன்ற* பெயரடைகளின் பின்னும் *வராத*, *படிக்காத* போன்ற பெயரெச்சங்களின் பின்னும் ஒற்று மிகாது.

நல்ல + பையன் = நல்ல பையன்

வராத + பணம் = வராத பணம்

6. *சின்ன* என்னும் பெயரடை ஏனைய பெயரடைகள் அல்லாமல் ஒற்று மிகுந்து வரும்.

சின்ன + கதவு = சின்னக் கதவு

7. மறைந்திருக்கும் *ம்* என்பது ஞ் என மாறி வரும்.

சின்ன(ம்) + சிறு = சின்னஞ்சிறு

8. *பல* என்பது இரட்டித்து *பலப்பல*, *பலபல*, *பற்பல* என வரலாம். *சில* என்பது இரட்டித்து *சிலசில*, *சிற்சில* என வரும்.

9. இல்லை உடம்படு மெய் *வ்* ஏற்று வரும்.

வர + இல்லை = வரவில்லை

10. தற்காலத் தமிழில் உடம்படுமெய் இல்லாமல் வரும்.

வந்த + இடத்தில் = வந்த இடத்தில்

11. *கடகட*, *வளவள* என்ற இரட்டைக் கிழவிகள் *என*, *என்று* என்பவைகளுடன் உடம்படுமெய் ஏற்று வரும்.

கடகடவன்று.

12. *என்ன* என்பது இரட்டித்து என்னென்ன என வரும்.

13. சில தொகைச் சொற்களில் ஒற்று மிகும்

பலா + கொட்டை = பலாக்கொட்டை

திருவிழா + கூட்டம் = திருவிழாக்கூட்டம்

14. *ஓடி* போன்ற வினை எச்சத்தின் முன் ஒற்று மிகும்.

ஓட்டி + போ = ஓடிப்போ

15. முதற்சொல் வினையடியாக இருந்தால் ஒற்று மிகாது.

சுடு + சோறு = சுடு சோறு

குடி + தண்ணீர் = குடி தண்ணீர்

16. சில கூட்டுச் சொற்கள் உருவாக்கத்தில் *ல்* என்பது *ற்* என மாறும். சிலவற்றில் மாறாது.

கல் + சிலை = கற்சிலை

பல் + பொடி = பற்பொடி

நாள் + காட்டி = நாட்காட்டி

17. *கு*, *சு*, *டு*, *ணு*, *து*, *பு*, *று*, *ரு*, *மு*, *வு* ஆகிய எழுத்துக்களில் முடியும் பெயர்ச்சொற்கள் கூட்டுச் சொற்களாக வருகையில் ஒற்று மிகும்.

தேக்கு + கட்டை = தேக்குக்கட்டை

18. இரண்டாம் வேற்றுமை உருபு *ஐ* பின் ஒற்று மிகும்.

பழத்தை + கடித்தான் = பழத்தைக் கடித்தான்

19. வேற்று மை உருபு *கு* பின் ஒற்று மிகும்.

அவனுக்கு + கொடு = அவனுக்குக் கொடு

20. வேற்றுமை உறவில் அல்லது அடையா வரும் *ம்* இறுதிப் பெயர்கள் மற்றொரு பெயருடன் இணையும் போது ம் கெட ஒற்று மிகும்.

மரம் + பெட்டி = மரப்பெட்டி

21. வேற்றுமை உறவில் அல்லது அடையாக வரும் பெயர்ச்சொற்கள் இணைகையில் ஒற்று மிகும்.

பேய் + காற்று = பேய்க்காற்று

தயிர் + சோறு = தயிர்ச்சோறு

22. சில கூட்டாக்கங்களில் *ல்* என்பது *ற்* ஆகவும் *ள்* என்பது *ட்* ஆகவும் மாறும். சிலவற்றில் மாறாது அல்லது விருப்பாக மாறவோ மாறாமலோ வரும்.

பல் + பொடி = பற்பொடி

முள் + செடி = முட்செடி

நாள் + காட்டி = நாட்காட்டி/ நாள் காட்டி.

3.6 தமிழில் பிழை ஆய்வு

தமிழில் தட்டச்சு செய்யப்படும் பிழைகள் விசைப் பிழையாலும் அறியாமலும் தவறான எழுத்துத் தேர்வாலும் ஏற்படலாம். தமிழ் பிழை ஆய்வு குறித்து நடராச பிள்ளை-விலா எழுதிய “பிழை ஆய்வு” என்ற நூலில் கூறப்பட்ட கருத்துக்களும் தமிழ் பிழை திருத்தி உருவாகத்திற்குத் துணை புரியும்.

3.6.1 உச்சரிப்பு ஒற்றுமை காரணமான பிழைகள்

தமிழ் *ல்*, *ள்*, *ழ்* என்ற எழுத்துக்கள் உச்சரிப்பு என்ற நிலையில் தமிழ் நன்றாக அறியாதவர்களுக்கும் தமிழை இரண்டாம் மொழியாகக் கற்பவர்களுக்கும் ஒன்றுபோல் தோன்றுவதால் ஒன்றிற்குப் பதிலாக ஒன்றை எழுத்துவது வழக்கம். எடுத்துக்காட்டாக *வாழை* என்பது *வாளை* என எழுதப்படலாம். இந்த இரு சொற்களும் அகராதியில் இருப்பதால் கணிணிக்கு சூழல் அறிவு தரப்பட்டால் தான் பிழை திருத்தம் செய்ய இயலும்.

எ.கா.

வ.எண்	பிழை	திருத்தம்
1	ஆளமாக	ஆழமாக
2	ஆள்கடல்	ஆழ்கடல்
3	அளைகடல்	அலைகடல்
4	பளம்	பழம்
5	தமிள்	தமிழ்

ந், *ன்*, *ண்* என்ற எழுத்துக்களும் உச்சரிப்பு என்ற நிலையில் தமிழ் நன்றாக அறியாதவர்களுக்கும் தமிழை இரண்டாம் மொழியாகக் கற்பவர்களுக்கும் ஒன்றுபோல் தோன்றுவதால் ஒன்றிற்குப் பதிலாக ஒன்றை எழுத்துவது வழக்கம். எடுத்துக்காட்டாக *நகர்* என்பது *னகர்* என்றோ *ணகர்* என்றோ எழுதப்படலாம்; *அண்ணன்* என்பது *அன்னன்* என்று எழுதப்படலாம். பிழையாய் எழுதப்பெற்றும் அகரதியில் உள்ள சொற்களைச் சூழல் அறிவு இல்லாமல் பிழை திருத்தம் செய்ய இயலாது.

ர், *ற்* என்ற எழுதத்துக்களும் உச்சரிப்பு என்ற நிலையில் தமிழ் நன்றாக அறியாதவர்களுக்கும் தமிழை இரண்டாம் மொழியாகக் கற்பவர்கும் ஒன்றுபோல் தோன்றுவதால் ஒன்றிற்குப் பதிலாக ஒன்றை எழுத்துவது வழக்கம். *அறம்* என்பது *அரம்* என எழுதப்படலாம். இங்கும் பிழையாய் உள்ள *அரம்* என்ற சொல்லும் அகராதியில் உள்ளதால் சூழல் அறிவு கணிப்பொறிக்குத் தரப்பட்டால் தான் பிழை திருத்தம் செய்ய இயலும்.

எ.கா.

வ.எண்	**பிழை**	**திருத்தம்**
1	வகுப்பு அரை	வகுப்பு அறை
2	கடல் கறை	கடல் கரை
3	குரைவாக	குறைவாக

3.6.2 ஒலியனியல் அமைப்பு அறியாததால் ஏற்படும் பிழைகள்

தமிழின் ஒலியன் வருகை முறை விதிகள் (phonotactic ruels) அறியாததன் காரணமாகப் பிழைகள் ஏற்படும். இத்தகைய பிழைகளும் எழுத்துப் பிழை திருத்தி உருவாக்கத்திற்கு இன்றியமையாத தரவாகும். ஒரு சில எடுத்துக்காட்டுகள் கீழே விளக்கப்பட்டுள்ளன.

ரகரத்திற்குப் பின் வல்லின எழுத்துக்கள் இரட்டித்து வரும் (எ.கா. *கவர்ச்சி*, *அவர்க்கு*); ஆனால் றகரத்திற்குப் பின் இவ்வாறு இரட்டுவதில்லை. இது சந்தி விதி அறியாமையால் ஏற்படும் பிழையாகும்.

வ.எண்	தவறு	சரி
1	முயற்ச்சி	முயற்சி
2	ஏற்ப்படுத்து	ஏற்படுத்து
3	கற்க்கிறேன்	கற்கிறேன்

இது போன்று *ட்*-க்கு பின்னரும் வல்லினம் மிகா.

வ.எண்	தவறு	சரி
1	காட்ச்சி	காட்சி
2	தட்ப்பம்	தட்பம்
3	உட்க்கார்	உட்கார்

ழ்-க்கு பின் வல்லெழுத்து மிகாது எழுதுவதும் இம்மாதிரியான விதி அறியாப் பிழைகளாகும்.

வ.எண்	தவறு	சரி
1	வாழ்கை	வாழ்க்கை
2	மகிழ்சி	மகிழ்ச்சி
3	வாழ்து	வாழ்த்து

3.6.3 சந்தி விதிகள் அறியாததால் ஏற்படும் பிழைகள்

சந்தி விதிகள் அறியாததால் பலர் தவறுகள் செய்வர். அத்தகைய பிழைகள் பொதுவாக எழுத்து ஆவணங்களிலும் தட்டச்சு செய்யப்பட்ட ஆவணங்களிலும்

பரவலாகக் காணப்படும். சந்தி விதிகள் அறிந்தவர்கள் கூடக் கவனக்குறைவு காரணமாக இத்தகைய பிழைகள் செய்வர். இத்தகைய பிழைகளையும் எழுத்துப் பிழை திருத்தி கண்டுபிடித்து நேர் செய்யவேண்டும். மாணவர்களின் விடைத்தாள்களைத் திருத்தும் போது இத்தகைய பிழைகள் மிகுந்து கிடப்பதை நாம் காண்கின்றோம். சரியாகப் பிழை திருத்தம் செய்யப்படாத ஆவணங்களிலும் கட்டுரைகளிலும் நூல்களிலும் இத்தகைய பிழைகள் மிகுந்து கிடப்பதைக் காணலாம். இத்தகைய பிழைகளும் கண்டுபிடிக்கப்பட்டு திருத்தப்படவேண்டும். சந்திப்பிழைகள் இருப்பது படிப்பவர்களின் கவனக்குவிப்பைச் சிதறச் செய்வதோடு தவறைப் படிப்பதால் ஏற்படும் மன உளைச்சலையும் ஏற்படுத்தும். மேற் சொன்ன தவறுகளைப் பட்டியலிட்டு ஒரு எழுத்துப் பிழைத் தரவுத்தொகுப்பை உருவாக்க இயலும். அத்தரவுத்தொகுப்பை எழுத்துப்பிழை திருத்தி உருவாக்கத் தரவாகப் பயன்படுத்தலாம்.

வ.எண்	**தவறு**	**சரி**
1	மலைக்கள்	மலைகள்
2	பொனை	பொன்னை
3	தூண்ணை	தூணை
4	கணை	கண்ணை
5	பலில்	பல்லில்
6	இந்தியாயில்	இந்தியாவில்
7	மாமாயிடம்	மாமாவிடம்
8	கைதொழில்	கைத்தொழில்
9	என் உடன்	என்னுடன்
10	என் இடம்	என்னிடம்
11	பாட புத்தகம்	பாடப்புத்தகம்

3.6.4 பேச்சுத் தமிழ் இடையீட்டுப் பிழைகள்

பேச்சுத் தமிழின் தாக்கம் காரணமாக எழுதும் மொழியில் பிழை ஏற்பட வாய்ப்புண்டு. பின்வருவன எடுத்துக்காட்டுகளாகும்.

வ.எண்	தவறு	சரி
1	தேச்சேன்	தேய்த்தேன்
2	படிச்சான்	படித்தான்
3	பாக்கணும்	பார்க்கவேண்டும்
4	போற	போகிற

3.6.5 தாய்மொழி இடையீடுப் பிழை

பயில் மொழியில் தாய் மொழியின் இடையீடு காரணமாக பயில் மொழிப் பயன்பாட்டின் எல்லா நிலைகளிலும் பிழைகள் ஏற்படும். குறிப்பாகச் தமிழை இரண்டாம் மொழியாகக் கற்பவர்கள் தமிழ் எழுத்துக்களுக்குப் பதிலாக வடமொழி எழுத்துக்களைப் பிழையாகப் பயன்படுத்துவர். பின்வரும் எடுத்துக்காட்டுகள் இந்நிலையைத் தெளிவு படுத்தும்.

வ.எண்	தவறு	சரி
1	ஸாதம்	சாதம்
2	ஸந்தோஷம்	சந்தோசம்
3	ஸாப்பாடு	சாப்பாடு
4	தஹரம்	தகரம்
5	மஹன்	மகன்

3.6.6 உருபனியல் பிழைகள்

பெரும்பாலும் உருபனியல் பிழைகள் ஒப்புமையாக்கத்தின் அடிப்படையிலேயே நிகழ்கின்றது.

எ.கா.

விழுந்தான் என்ற ஆக்கப்படி தொழுந்தான் என கால ஒட்டு தவற எழுதுவது.

தேடு - தேட என்ற ஆக்கப்படி கொடு என்பதை கொடுக்க என எழுதுவதற்குப் பதிலாகக் கொட என எழுதுவது.

3.7 முடிவுரை

இவ்வியலில் தமிழ் மொழியின் எழுத்தியல் மற்றும் ஒலியனியல் அமைப்பு பற்றி விரிவாகப் பார்த்தோம். தமிழ் மொழியின் புணர்ச்சி விதிகள் பற்றி தெளிவு படுத்தப்பட்டது. தமிழின் உருபனியல் பற்றி விரிவாகக் கூறப்பட்டது. தமிழில் செய்யப்படும் பிழைகளின் ஆய்வுபற்றிச் சுருக்கமாகக் கூறப்பட்டது. இவையாவும் தமிழ் மொழிக்கு உருவக்க இருக்கும் எழுத்துப் பிழை மற்றும் இலக்கணப் பிழை திருத்திக்கு உறுதுணையாக அமையும்.

இயல் 4

தமிழில் எழுத்துப்பிழை மற்றும் இலக்கணப்பிழை திருத்தி உருவாக்கம்

4.0 முன்னுரை

இங்கு தமிழுக்கென்று உருவாக்கப்பட்ட எழத்துப் பிழை மற்றும் இலக்கணப் பிழை திருத்திகளின் சுருக்கமான தகவல்கள் தரப்படும். மேலும் இவ்வாய்வின் விளைவாக உருவாக்கப்பட்ட எழுத்துப் பிழை மற்றும் இலக்கணப் பிழை திருத்தியின் மாதிரிகள் விளக்கப்படும்.

4.1 தமிழுக்கென்று உருவாக்கப்பட்ட எழுத்துப் பிழை திருத்திகளின் சுருக்கம்

தமிழுக்கென்று எழுத்துப்பிழை திருத்தி உவாக்கும் முயற்சிகள் பல மேற்கொள்ளப்பட்டுள்ளன. சில சொல்லாய்விகளில் இத்தகைய எழுத்துப்பிழை திருத்திகள் பயன்படுத்தப்படுகின்றன. முழுமையான எழுத்துப் பிழை திருத்தி உருவாக்கம் இப்போதும் முயற்சிக்கப்பட்டுவருகின்றது. இதுவரை உருவாக்கப்பட்ட எழுத்துப் பிழை திருத்திகள் பற்றி இங்குச் சுருக்கமாக விளக்கப்பட்டுள்ளது.

4.1.1 தமிழ்ப் பல்கலைக்கழகப் பிழை திருத்தி

தமிழ்ப் பல்கலைக் கழக மொழியியல் துறையில் இராசேந்திரன் அவர்களால் 2000-2001 காலகட்டத்தில் தமிழுக்கு எழுத்துப் பிழை திருத்தியும் இலக்கணப்பிழை திருத்தியும் உருவாக்கும் முயற்சி மேற்கொள்ளப்பட்டது. பல்கலைக்கழக மானியக் குழுவின் நிதிநல்கையால் இது மேற்கொள்ளப்பட்டது. இது உருபனியல் பகுப்பாய்வி அடிப்படையில் அமைந்திருந்தது. உருபனியல் பகுப்பாய்வி அனுசாரகா திட்டம் அடிப்படையில் அமைந்திருந்தது. இது ஒரு அடுக்குமுறை அணுகுமுறையாகும் (paradigmatic approach). இதில் வினைகளும் பெயர்களும் அவைகளின் ஒலியனியல்-உருபனியல் நடத்தை அடிப்படையில் குழுக்களாகப் பிரிக்கப்பட்டு அதன் அடிப்படையில் வினைவடிவங்களும் பெயர்வடிவங்களும் ஆயப்பட்டன. தவறான உருபனியல் பகுப்பாய்வு விளைந்தால் அச்சொல் தவறு என்று சுட்டிக்காட்டப்படும். சரியன உருபனியல் பகுப்பாய்வு விளைந்தால் சொற்கள் சரியானவை என்று சுட்டிக்காட்டப்படும். இந்த அளவில் தான் இவ்வாய்வு அமைந்தது. இதற்கு மேல் இவ்வாய்வு மேற்கொள்ளப்படவில்லை. இருப்பினும் தமிழின் ஒலியனியல் அமைப்பும் உருபனியல் அமைப்பும் ஆயப்பட்டு தகவல்கள் திரட்டப்பட்டன. ஒலியன் வருகைமுறை

விதிகள், உருபன் வருகை முறை விதிகள், சொற்கள் மற்றும் தொடர்கள் வருகை முறை விதிகள் யாவும் திரட்டப்பட்டு கணினி மொழியியல் ஆய்வுக்காக பட்டியலிடப்பட்டன. சொற்களின் வகைப்பாடு அடையாளப்படுத்தப்பட்ட ஒரு மின் அகராதியும் உருவாக்கப்பட்டது.

4.1.2. தனபாலன் மற்றும் பிறரின் தமிழ் எழுத்துப்பிழை திருத்தி

தனபாலன் மற்றும் பிறர் தமிழுக்கு என்று உருவாக்கிய எழுத்துப் பிழை திருத்தி (Dhanabalan, Ranjani Parthasarathi and Geetha) அகராதிகளைப் பயன்படுத்துகின்றது. உருபனியல் மற்றும் தொடரியல் தகவல்களைக் கொண்ட அகராதிகள் சொல் ஆய்விகளில் ஒருங்கிணைக்க இயலும் தமிழ் எழுத்துப்பிழை திருத்தி உருவாக்கவும் அதிக கலவைத் தன்மையான இயற்கைமொழி ஆய்வுப் பயன்பாடுகளில் பயன்படுத்த இயலும் உருபனியல் மற்றும் தொடரியல் ஆய்விகளின் உருவாக்கத்திற்கும் தேவைப்படுகின்றது.

4.1.2.1. தமிழ்ப் பிழைதிருத்தத்தின் சிக்கல்கள்

இவர்கள் பிழை திருத்தத்தின் முக்கியச் சிக்கல்களாகப் பின்வருவனவற்றைப் பட்டியலிடுகின்றனர்.

1. வேற்றுமை உருபுகளைக் கையாளுதல்: ஆங்கிலத்தில் தனிச் சொல்லாக வரும் முன்னுருபுகள் தமிழில் வேற்றுமை உருபுகளாகவும் (பின்னொட்டுகளாகவும்) முன்னுருபுகளாகவும் வரும்.

   எ.கா. ராமன் கிருஷ்ணனைப் பற்றிச் சொல்கின்றான்.

2. வேற்றுமை உருபுகளால் வேர்ச்சொல்லைப் பெயராகவோ வினையாகவோ வகைப்படுத்துதல்.

3. குற்றுயிர்-நெட்டுயிர் பிழைகளைக் கையாளுதல்.

   கிணறு - கேணறு

4. இரு சொற்களுக்கு இடையில் வரும் புணர்ச்சி விதிகளைக் கையாளுதல்.

   அவனுக்குப் பிறகு

4.1.2.2 தமிழ்ப் பிழைதிருத்தியின் தொகுதிகள்

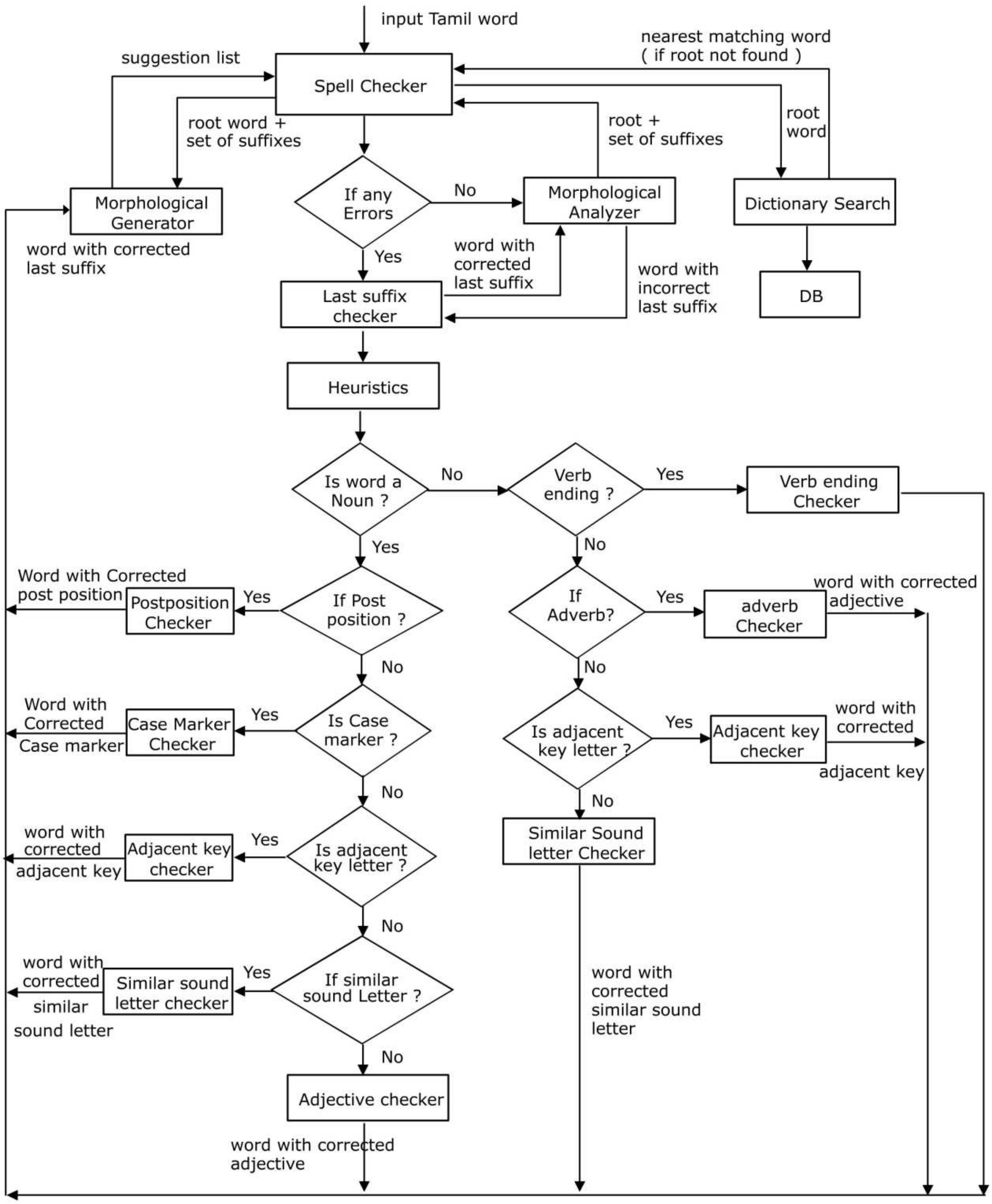
தமிழ் பிழை திருத்தியின் பொதுவான அமைப்பு ஒழுக்குப் படத்தில் தரப்பட்டுள்ளது. தொடக்கத்தில் பிழைதிருத்தி ஆவணத்திலிருந்து எடுக்கப்பட்ட சொற்களை ஒரு வேளைக்கு ஒன்று என்று படிக்கின்றது. அகராதி பின்னர் கிடைத்த சொல்லை பரிசோதிக்கின்றது. சொல் அகராதியில் காணப்பட்டால் அது ஒரு சரியான சொல்லாகக் கருதப்படும். சொல் அகரதியில் காணப்படாவிட்டால் அது பிழைதிருத்தத்திற்கு அனுப்பப்படும். பிழை திருத்தி பின்வரும் பகுதிகளைக் கொண்டது:

- பனுவல் பகுத்தாய்தல்
- எழுத்துக் கூட்டல் பரிசோதனை மற்றும் திருத்தல்
- பரிந்துரைப் பட்டியல் உருவாக்கம்

இப்பகுதிகளுக்கு உதவப் பிழை திருத்தி பின்வருவனவற்றைப பயன்படுத்தும்.

- தரப்பட்ட சொல்லை ஆயும் உருபனியல் ஆய்வி
- பரிந்துரைகளை உருவாக்கும் உருபனியல் உருவாக்கி

பனுவல் பகுத்தாய்தல் உள்ளீடு செய்யப்பட்ட ஆவணத்தில் பிழை திருத்துவதற்குப் பொருத்தமான ஒரு சொல்லை அதைச் சுற்றியுள்ள பனுவலிலிருந்து தனியாகப் பிரிக்கின்றது. கண்டுபிடிக்கப்பட்ட சொல் பிழையானது என்றால் அது பின்னர் எழுத்துப்பிழை திருத்தப் பகுதிக்கும் பரிந்துரைகள் உருவாக்கப் பகுதிக்கும் அனுப்பப்படும். ஆவணங்களைச் சொற்களின் பட்டியலாகப் பகுத்தாய்ந்த பின்னர் ஒவ்வொரு சொல்லும் உருபனியல் பகுப்பாய்விக்கு (morphological analyzer) அனுப்பப்படும். உருபனியல் பகுப்பாய்வி பின்னொட்டைப் பிரிக்க முயலும். அது சரியான சொற்களை மட்டும் பகுத்தாயும் படிக்குத் திட்டமிடப்பட்டுள்ளது. எழுத்துப்பிழை காரணமாக அது பின்னொட்டைப் பிரிக்க இயலவில்லை என்றால் அது அச்சொல்லை எழுத்துக் கூட்டல் பரீசோதனை மற்றும் திருத்தப் பகுதிக்கு அனுப்பும். சொல்லை வேர்களாகவும் பின்னொட்டுக்களாகவும் பிரித்த பின்னர் அகராதியுடன் ஒப்பிடப்பட்டு அதன் செல்லுபடி சரிபார்க்கப்படும். அச்சொல் அகராதியில் இல்லை என்றால் அண்மை அடிப்படையில் பொருந்துகிற பரிந்துரைகள் தெரிந்தெடுக்கப்படும்.


input Tamil word
suggestion list
Spell Checker
nearest matching word (if root not found)
root word + set of suffixes
root + set of suffixes
root word
If any Errors
No
Yes
Morphological Generator
Morphological Analyzer
Dictionary Search
DB
word with corrected last suffix
word with corrected last suffix
word with incorrect last suffix
Last suffix checker
Heuristics
Is word a Noun ?
Verb ending ?
Verb ending Checker
If Post position ?
Postposition Checker
Word with Corrected post position
If Adverb?
adverb Checker
word with corrected adjective
Is Case marker ?
Case Marker Checker
Word with Corrected Case marker
Is adjacent key letter ?
Adjacent key checker
word with corrected adjacent key
Similar Sound letter Checker
If similar sound Letter ?
Similar sound letter checker
word with corrected similar sound letter
Adjective checker
word with corrected adjective

4.1.2..2.1 பெயர்களைச் சரிபார்த்தல்

தமிழில் பெயர் ஒரு சொல்வகையாகும். இது வேற்றுமை உருபுகளையும் பன்மை உருபுகளையும் பின்னுருபுகளையும் குறைசொல் (clitics) பின்னொட்டுக்களையும் எடுக்க வல்லன. பெயர் சொல்பகுதிகளை ஓருபன் சொற்கள் (எ.கா. *கல்*, *மலர்*) என்றும் பல்லுருபன் சொற்கள் (எ.கா. *கல்வி*, *காட்சி*) என்றும் பகுக்க இயலும். ஓருபன் சொற்களை வேர்களாகவும் ஒட்டுக்களாகவும் பிரிக்க இயலாது; ஆனால் பல்லுருபன் சொற்களை வேர்களாகவும் ஒட்டுக்களாகவும் பிரிக்க இயலும். வேற்றுமை உருபுகள் பெயர்த்தொடருக்கும் வினைப் பயனிலைக்கும் இடையே உள்ள தொடரியல் மற்றும் பொருண்மை உறவுகளை வெளிப்படுத்தும். தமிழில் பெரும்பாலான வேற்றுமைகள் பின்னொட்டுக்களாலும் சில வேற்றுமை பின்னொட்டுக்களுடன் இணைந்த பின்னுருபுகளாலும் மெய்ப்படுத்தம் செய்யப்படும். இங்கு பத்து வேற்றுமைகள் கருத்தில் கொள்ளப்பட்டுள்ளன; அவற்றின் பட்டியல் பின்னொட்டுகளுடன் அட்டவணயில் தரப்பட்டுள்ளது.

வ.எண்	வேற்றுமை	பின்னொட்டு
1	எழுவாய் வேற்றுமை	0
2	செயப்படுபொருள் வேற்றுமை	ஐ
3	கொடை வேற்றுமை	கு
4	பயன் வேற்றுமை	உக்காக
5	கருவி வேற்றுமை	ஆல்
6	உடனிகழ்ச்சி வேற்றுமை	ஆல்
7	இட வேற்றுமை	இல்
8	நீங்கல் வேற்றுமை	இல்லிருந்து, இடமிருந்து
9	கிழமை வேற்றுமை	அது

பிழைதிருத்தச் செயல்பாடு வேற்றுமை உருபு சரிபார்த்தல், பின்னுருபு சரிபார்த்தல், பெயரடை சரிபார்த்தல் என்று மேலும் பகுக்கப்படும். *இல்*, *ஆல்*, *இடம்*, *உடன்*, *ஆக*, *கல்*, *உக்கு*, *ஓடு*, *இருந்து*, *உடைய*, *அது* போன்ற வேறுபட்ட வேற்றுமை உருபுகள் உருபனியல் பகுப்பாய்வியால் பிரித்தெடுக்கப்படுகின்றன. இக்கட்டத்திற்குப் பின்னர் இறுதி ஒட்டின் செல்லுபடியைத் தீர்மானிக்க ஒரு சரிபார்ப்பு நடத்தப்பெறும். பின்வருவன பெயர்களில் வேற்றுமை உருபுகளில் விளையும் சில பிழைகளாகும்.

எ.கா.

மரத்தில் என்பதற்குப் பதிலாக *மரத்தல்* என வருதல்

அவனிடம் என்பதற்குப் பதிலாக *அவனேடம்* என வருதல்

அவனுக்காக என்பதற்குப் பதிலாக *அவனுக்க* என வருதல்

மரங்கள் என்பதற்குப் பதிலாக *மரங்காள்* என வருதல்

மரத்திலிருந்து என்பதற்குப் பதிலாக *மரத்திலருந்து* என வருதல்

மரத்துக்கு என்பதற்குப் பதிலாக *மரத்தக்கு* என வருதல்

மரத்திலருந்து என்ற சொல்லை எழுத்துப்பிழை சரிபார்க்க எடுத்துக்கொள்ளவும். இச்சொல் பகுப்பாய்வுக்காக உருபனியல் பகுப்பாய்விக்கு அனுப்பப்படும். சொல்லின் பின்னொட்டு *இலருந்து* என்று இருப்பதன் காரணமாக உருபனியல் பகுப்பாய்வியால் இச்சொல்லைப் பிரிக்க இயலவில்லை. எழுத்துக்கூட்டல் சரிபார்ப்பும் பிழைதிருத்தக் கட்டமும் இறுதி பின்னொட்டை (*இலருந்து* என்பதை) *இலிருந்து* எனத் திருத்தி சொல்லைப் பிரிப்பதற்காக உருபனியல் பகுப்பாய்விக்கு அனுப்பும். இறுதியில் அது வேர்ச்சொல்லையும் பின்னொட்டையும் பரிந்துரை உருவாக்கக் கட்டத்திற்கு அனுப்பும்.

4.1.2.2.2 வினைச் சொற்களைச் சரிபார்த்தல்

கால ஒட்டுக்களை ஏற்கவல்ல சொல்வகைப்பாடு வினையாகும். வினைச் சொல் வினையடையால் அல்லது வினையெச்சத் தொடரால் அடைசெய்யப்படலாம்; வினை இடம்-எண்-பால் ஒட்டுக்களால் திரிபுறலாம் மற்றும் பெயரெச்சமாகவோ வினை

எச்சமாகவோ மாற்றப்படலாம். வினைகள் இரு குழுமங்களாகப் பகுக்கப்படும்: ஒன்று முற்று வடிவம்; மற்றொன்று எச்சவடிவம்.

4.1.2.2.2.1 முற்றுவடிவம்

வினையின் முற்றுவடிவம் தனிநிலை வடிவாகவோ கூட்டுநிலை வடிவாகவோ இருக்கலாம். தனிநிலை முற்று வடிவங்கள் அமைப்பு வகுப்புகளின் படி மூன்று குழுமங்களாகப் பகுக்கப்படும்.

1. பின்னொட்டுகள் எடுக்காத வினைகள். எ.கா. *சரி*, *தனி*.
2. இடம்-எண்-பால் ஒட்டு ஏற்காத காலமில்லா வினைகள். எ.கா. *இல்லை*, *உண்டு*
3. இடம்-எண்-பால் ஒட்டுக்களால் திரிபுறும் மற்றும பிற ஒட்டுக்களையும் ஏற்கும் வினைகள். எ.கா. *சாப்பிடு*

வினையடிகள் அவைகள் ஏற்கும் கால ஒட்டுகள் அடிப்படையயில் வகைப்படுத்தப் படவேண்டும்.

தனிநிலை முற்றுவினைகள்

கால ஒட்டுகளையும் இடம்-எண்-பால் ஒட்டுக்களையும் ஏற்கும் வினைகள் தனிநிலை வினைகள் என அழைக்கப்பெறும். இதில் வேர்களைத் தொடர்ந்து கால ஒட்டுகளும் அதனைத் தொடர்ந்து இடம்-எண்-பால் ஒட்டுக்களும் சேர்க்கப்படும். மனித எழுவாய்களுக்கும் மனிதரல்லாத எழுவாய்க்களுக்கும் கால ஒட்டுகள் எப்பொழுதும் ஒன்றாக இருப்தில்லை. மேலும் மனிதரல்லாத ஒருமை எழுவாய்க்கும் மனிதரல்லாத பன்மை எழுவாய்க்கும் எப்பொழுதும் கால ஒட்டுகள் ஒன்றாக இருப்பதில்லை. இடம்-எண்-பால் ஒட்டுக்கள் எழுவாயுடன் உடன்பாடு கொண்டிருக்கும். கால ஒட்டுக்குப் பதிலாக எதிர்மறை ஒட்டான ஆ என்பதையும் வினைவேருடன் நேரடியாகச் சேர்த்து எதிர்மறை வினைவடிவைப் பெறவியலும்.

எ.கா.

செய்-கிற்-ஆன் (நிகழ்காலம்)

செய்-த்-ஆன் (இறந்தகாலம்)

செய்-வ்.ஆன் (எதிர்காலம்)

செய்ய்-ஆ-து (எதிர்மறை வடிவம்)

கூட்டுநிலை முற்று வினைகள்

கூடுநிலை வினை வினையடி, வினையாறுவகை, வினைப்பாடு, வினைநோக்கு, காலம், இடம்-எண்-பால் மற்றும் குறைச்சொல் இவற்றைக் கொண்டிருக்கும். இவற்றில் வினையடியைத் தொடரும் ஒட்டுகள் விருப்பாகும். கூட்டுநிலை வினைகள் ஒன்றோ அதற்கு மேலோ துணைவினைகளைக் கொண்டிருக்கும். பொதுவாக ஒருவகைத் துணைவினனை அதேபோன்ற துணைவினையைத் தொடராது.

4.1.2.2.2.2 வினையின் எச்ச வடிவங்கள்

வினை எச்ச வடிவங்களைப் பெயரெச்ச வடிவம், வினையெச்ச வடிவம் என்று வகைப்படுத்தலாம். பெயரெச்ச வடிவைக் கால ஒட்டுடன் *அ* என்ற ஒட்டைச் சேர்த்துப் பெறலாம்.

எ.கா.

படி-த்த்-அ > படித்த

பதி-க்கின்ற்-அ > படிக்கின்ற

படி-உம் > படிக்கும்

வினையுடன் *அ* அல்லது *து* அல்லது *ஆல்* என்ற ஒட்டை நேரடியாகச் சேர்த்து வினையெச்ச வடிவைப் பெறலாம்.

எ.கா.

போ-க > போக (செய்ய வினையெச்சம்)

படி-க்க > படிக்க (செய்ய வினைஎச்சம்)

பார்-த்து > பார்த்து (செய்து வினையெச்சம்)

செய்-த்-ஆல் > செய்தால் (நிபந்தனை வினையெச்சம்)

படி-த்-ஆல் > படித்தால் (நிபந்தனை வினையெச்சம்)

வினை சரிபார்ப்பு இடம்-எண்-பால் சரிபார்ப்பு, கால ஒட்டு சரிபார்ப்பு என மேலும் வகைபடுத்தப்படும். வினையின் வேறுபட்ட ஒட்டுகள் உருபனியல் ஆய்வியால் பிரித்தெடுக்கப்படும். கால ஒட்டுகள் *கிற்*, *கின்று*, *ஆனின்று*, *ப்*, *வ்*, *த்*, *ட்*, *ற்* என்பனவற்றை உட்படுத்தும். இவை வினைகளை நிகழ்காலம் அல்லது இறந்தகாலம் அல்லது எதிர்காலம் என்று அடையாளப்படுத்தும். இடம்-எண்-பால் ஒட்டுகள் *அன்*, *ஆன்*, *அள்*, *ஆள்*, *அர்*, *ஆர்*, *து*, *அ*, *வை*, *கள்* என்பனவற்றை உட்படுத்தும். வினையின் ஒட்டுகளில் சொல் பிழையாக எழுதப்பட்டால் பிழை திருத்தப்பட்டு சாத்தியமான பரிந்துரைகள் உருவாக்கப்படும்.

எ.கா.

செய்கிறார்கள் என்பதற்குப் பதிலாகச் *செய்கிறர்கள்* என வருதல்
செய்தது என்பதற்குப் பதிலாகச் *செய்தத* என வருதல்
செய்வான் என்பதற்குப் பதிலாகச் *செய்வன்* என வருதல்

தோன்றுகின்றன் என்ற சொல்லைக் கருத்தில் கொள்ளவும். இச்சொல் பகுத்தாயப்படுவதற்காக உருபனியல் பகுப்பாய்வியிடம் அனுப்பப்படுகின்றது. எழுத்துக்கூட்டல் சரிபார்ப்பு, திருத்தம் என்ற கட்டங்கள் *தோன்றுகின்றன்* என்ற சொல்லின் இறுதி ஒட்டான *அன்* என்பதை *ஆன்* என திருத்துகின்றது. பகுப்பாய்வி வேர்ச் சொல்லையும் பின்னொட்டுக்களையும் பரிந்துரை உருவாகக்கட்டத்திற்கு அனுப்புகின்றது. உருபனியல் பகுப்பாய்வியால் அனுப்பப்பட்ட வேர்ச் சொல்லையும் பின்னொட்டுகளையும் பரிந்துரை உருவக்கக் கட்டம் பரிந்துரை உருவாக்கத்திற்காக உருபனியல் உருவாக்கியிடம் அனுப்புகின்றது.

4.1.2.2.2.3 பின்னுருபுகளைச் சரிபார்த்தல்

வேற்றுமை உருபின் செயல்பாட்டைச் செய்யும் சொல் பின்னுருபு என அழைக்கப்படும். பொதுவாக பின்னுருபு வேற்றுமை பின்னொட்டிற்குப் பின்னர் வரும். சில இடங்களில் வேற்றுமை பின்னொட்டு விருப்பாக வரும். பின்னுருபுகள் அவற்றின் முன்வரும் வேற்றுமை உருபுகளின் அடிப்படையில் ஐந்து வகுப்புகளாக வகைபடுத்தப்படுகின்றன.

செயப்படுபொருள் வேற்றுமையின் பின்னர் வரும் பின்னொட்டுகள்:

விட, *போல*, *கொண்டு*, *பற்றி*, *குறித்து*, *சுற்றி*, *விட்டு*, *தவிர*, *முன்னிட்டு*, *வேண்டி*, *ஒட்டி*, *பொறுத்து*, *பொறுத்தவரை*, …

உடைமை வேற்றுமையின் பின்னர் வரும் பின்னொட்டுகள்:

மீது, *மேல்*, *வழியாக*, *மூலமாக*, *வாயிலாக*, *கீழ்*, …

இட வேற்றுமையின் பின்னர் வரும் பின்னொட்டுகள்:

இருந்து, *உள்ள*

கொடை வேற்றுமையின் பின்னர் வரும் பின்னொட்டுகள்:

ஆக, *என்று*, *முன்*, *பின்*, *உரிய*, *உடைய*, *மத்தியில்*, *எதிரில்*, *அருகில்*, *பதில்*, *தகுந்த*, *மாறாக*, *நேராக*, …

உள்ளீட்டுச் சொல் ஒரு பின்னொட்டாக இருந்தால் அது சரியா தவறா என்று சரிபார்க்கும். சொல் தவறு என்றால் நேராக்கப்படும் மற்றும் பரிந்துரைகள்/ஆலோசனைகள் உருவாக்கப்படும்.

4.1.2.2.2.4. ஒன்றுபோல் உச்சரிக்கப்படும் எழுத்துக்கள்

ஒன்றுபோல் உச்சரிக்கப்படும் எழுத்து சொல்லின் எழுத்துமுறையில் தவறை விளைவிக்கக்கூடும். எடுத்துக்காட்டாகத் *தாளம்* என்ற சொலை எடுத்துக்கொள்வும். *ள* என்ற எழுத்து *ல* எனத் தவறாக (தாலம்) எழுதப்படலாம். தமிழில் மூன்று எழுத்துக்களுக்கு அவற்றைப் போலவே உச்சரிக்கப்படும் எழுத்துக்கள் உள்ளன. தவறான சொல்லுக்குச் சாத்தியமான ஒன்றுபோல் உச்சரிக்கப்படும் எழத்துக்களைப் பரிசோதித்து அகராதியில் உள்ள பரிந்துரைகள் பயன்படுத்துவோருக்குக் காட்டப்படும். தமிழில் ஒன்றுபோல் உச்சரிக்கப்படுகின்ற சாத்தியமான எழுத்துக்கள் பின்வருவனவாகும்.

ல, ள, ழ

ர, ற

ண, ந, ன

4.1.2.2.2.5 அண்மை விசைப் பிழைகள்

தட்டச்சு செய்யும் போது பயன்படுத்துவோர் சரியான ஒன்றிற்குப் பதிலாகத் தவறான ஒன்றைத் தட்டச்சு செய்யலாம். எனவே தவறாகத் தட்டச்சு செய்யப்பட்ட

எழுத்தின் அடுத்துவரும் விசைகளைக் கருத்தில் கொள்வது முக்கியமான பங்களிப்பு செய்கின்றது. தவறாகத் தட்டச்சு செய்யப்பட்ட எழுத்து ஏதாவது அண்மை விசையுடன் பொருந்தினால் தவறான எழுத்துக்குப் பதிலாகச் சரியானதை இடம்பெயர்க்கவும். அதன்பின்னர் திருத்தப்பட்ட சொல்லைச் சாத்தியமான பரிந்துரைகளை உருவாக்க அகராதியில் அல்லது வேற்றுமை விதிகளால் சரிபார்க்கவும். அண்மை விசைத் திருத்தங்கள் தமிழ் 99 என்ற பண்பட்ட விசைப்பலகை அடிப்படையில் அமையும்.

எடுத்துக்காட்டாக '*அவனுடைய*' என்ற சொல் *அவனுடஎய* அல்லது *அவனுடோய* என்று தவறாகத் தட்டச்சு செய்யப்படக்கூடும். இவைகள் *டை* என்ற எழுத்திற்குச் சாத்தியமான அண்மை விசைத் தவறுகளாகும். எல்லா சாத்தியமான அண்மை விசைகளும் இடம்பெயர்க்கப்பட்டு அகராதியில் சரிபார்க்கப்பட்டு தவறு திருத்தம் செயல்படுகின்றது.

எ.கா. அ - ஆ, ஏ, எ, அஔவ்

க - ள, ற, ப, ல ங

எ அ, ஆ, ஏ, உ, அஔவ், ஓ

ப ற, ந, க, ம, ல, ர

வேர்சொல்லில் உள்ள தவறுகளின் திருத்தம் முடிவுற்றால் தொடர்புள்ள ஒட்டுக்களுடன் திருத்தப்பட்டவைகள் பரிந்துரை உருவக்க நிலைக்கு அனுப்பப்படும்.

4.1.2.2.2.6 எழுத்து நிலைநிறுத்து விதிகள்

ஒரு தமிழ்ச் சொல் 12 உயிர்களில் ஒன்றாலோ 10 உயிர்மெய்களில் (க,ச, த, ந, ப, ம, ய, வ, ய, ஞ) ஒன்றாலோ தொடங்கப்படவேண்டும். இதுபோல் தமிழ்ச் சொல் 12 உயிர்களில் ஒன்றாலோ 9 மெய்களில் (ண், ம், ன், ய், ர், ல், ழ், ள், ன்) ஒன்றாலோ இறுதியுறவேண்டும். சொல் இந்த குழுமத்திலிருந்து தொடங்காவிட்டால், பிழைதிருத்தி சாத்தியமான திருத்தங்களைச் செய்யும்; அகராதியிலிருந்து அண்மையிலுள்ள பொருத்தமான வேர்சொற்களைப் பெறும். இது போல் ஒரு சொல் முன்னர் கூறிய எழுத்துகளால் இறுதியுறாவிட்டால் வேற்றுமை விதிகளைப் பயன்படுத்தி பிழைதிருத்தி அதைத் திருத்தும்.

4.1.2.2.2.7அகராதியைப் பார்த்தல்

மற்றொரு முக்கியமான விசயம் அகராதியின் வடிவம், திரிபு மற்றும் ஆக்கத்தின் சிக்கல், அகராதி கோப்பு ஒழுங்கமைப்பு, அகராதி பிரிப்பு, சொல் பெறும் உபாயம்/தொழில்நுட்பம் மற்றும் பிறவற்றுடன் தொடர்புடைய கணினி அகராதி. இங்குப் பயன்படுத்தப்படும் தமிழ் அகராதி பெயர், வினை, பெயரடை, வினையடை, குறைச்சொற்கள் என்ற எல்லா வேர்ச்சொற்களையும் கொண்டிருக்கும். வேர்ச்சொல்லில் ஏதாவது பிழை இருந்தால் சொல்லின் திருத்தம் செய்யப்படும் மற்றும் அண்மைப்படும் பொருத்தமான அகராதி சொற்கள் சேகரிக்கப்படும். இந்த உருவாக்கம் பரிந்துரை உருவாக்க நிலைக்கு (suggestion generation phase) அனுப்பப்படும்.

4.1.2.2.2.8 பரிந்துரை உருவாக்கம்

பிழைதிருத்தி எல்லாச் சாத்தியமான பரிந்துரைகளையும் உருவாக்க உருபனியல் உருவாக்கியைப் பயன்படுத்தும். தமிழுக்குத் திட்டவரைவு செய்யப்பட்ட உருபனியல் உருவாக்கி பெயர்கள், வினைகள், பின்னுருபுகள், பெயரடைகள், வினையடைகள் மற்றும் பிற தொடரியல் வகைப்பாடுகளைத் தனித்தனியாகக் கையாள்வது அவசியாமாகும்; ஏனென்றால் இந்த ஒவ்வொரு தொடரியல் வகைப்பாட்டிற்கும் உருபனியல் கூறுகளின் சேர்க்கை வேறுபட்ட தகவல் வகை அடிப்படையிலானது.

வேர்களிலிருந்து பெயர்த் திரிபு ஆக்கங்களை உருவாக்கும் போது பன்மை ஒட்டின் வடிவைத் தீர்மானிக்கும் மொழியியல் விதிகள் கருத்தில் கொள்ளப்படவேண்டும். பெயர்களுக்கு வேற்றுமை உருபுகளை இணைப்பது உருபனியல் உருவாக்கியின் முக்கியமான வேலையாகும். மேலும் பெயர்கள் வேற்றுமை உருபுகளைத் திரிபடியாகத்தான் ஏற்க இயலும் என்பதையும் கருத்தில் கொள்ளவேண்டும். ஒலிநிரவல் ஒட்டுகள் சில நேர்வுகளில் திரிபடிசார் ஒட்டுகளுடனோ (oblique suffix) பன்மை ஒட்டுகளுடனோ இணைந்து வரும். பெயர் வேர்களுடன் முன்னர் சொன்ன ஒட்டுகளின் சேர்க்கையின் போது புணர்ச்சி விதிகள் கருத்தில் கொள்ளப்படவேண்டும்.

தமிழ் வினையின் உருபனியல் அமைப்பு இடம், பால், எண் போன்றவற்றை உருப்பபடுத்தம் செய்யும் ஒட்டுகளைக் கையாளுவதாலும் வினையாற்றுவகை, வினைநோக்கு, காரணவாக்கம், வினைப்பாங்கு போன்றவற்றைக் காட்டும் துணைவினைகளுடன் இணைவதாலும் அதிக அளவில் கலவைநிலை அடைந்துள்ளது.

வினையை உருபனியல் அடிப்படையில் உருவாக்கும் போது தெரிந்தெடுக்கப்பட்ட எழுவாயின் பால், எண், இடம் என்பனவற்றிகு ஏற்ப பொருத்தமான உடன்பாட்டு விகுதியைத் (வினைமுற்று விகுதியை) தெரிந்தெடுக்க வேண்டியது தேவையாகும்.

பிழை திருத்தியின் செயல்திறன் வழுநிலை பரிந்துரையாகத் தரப்படும் சரியான பரிந்துரை அடிப்படையில் அமையும். இந்நேர்வுகளில் பயன்படுத்துபவர் வழுநிலை பரிந்துரையை உறுதி செய்வது அவசியம்; பின்னர் அடுத்த தவறுக்கு முன்னேறலாம். இல்லையெனில் பயன்படுத்துபவர் பரிந்துரைகளின் பட்டியலைப் பார்த்து ஒன்றை தேர்ந்தெடுக்கவேண்டும். எல்லா பரிந்துரைகளும் அகராதியிலுள்ள சொற்களுடன் ஒப்பிடப்படவேண்டும். அகராதியிலுள்ள பரிந்துரைகள் பயன்படுத்துபவருக்குக் காட்டப்படும்.

4.1.2.3 முடிவுரை

தமிழுக்கான பிழைதிருத்தி தட்டச்சு செய்யும் போது ஏற்படும் பெருப்பாலான தவறுகளை அடையாளம் காணப் பயனாளிக்கு உதவுகின்றது. இந்தப் பிழை திருத்தியில் நடைமுறை செய்யப்பட்ட செயல்பாடுகள் பெயர்களின் வேற்றுமை உருபுகளையும் பின்னொட்டுகளையும் பிழை திருத்துதல், பெயர்களுக்குப் பெயரடைகளைத் திருத்துதல், வினைகளின் வினைமுற்று விகுதிகளைப் பிழைதிருத்துதல், வினையடை திருத்தம் மற்றும் அண்மை விசை பிழை திருத்தம் என்பனவாகும். தமிழ் பிழைதிருத்தியின் பயன்பாடு சொல் ஆய்விகள், தேடல் இயந்திரங்கள், தகவல் இறுப்பு மற்றும் தகவல் பிரித்தெடுப்பு ஒழுங்குமுறைகள், இயந்திர மொழிபெயர்ப்பு ஒழுங்குமுறைகள் என்பனவாகும்.

4.1.3 விஜய் சங்கர் ராமின் தமிழ் பிழைதிருத்தி உருவாக்குதல்

தமிழ் பிழை திருத்தி உருவக்குதல் குறித்து விஜய் சுந்தர் ராம் (Vija Sundar Ram 2012) தமது ஆய்வுக்கட்டுரையில் விளக்குகின்றார். பிழை திருத்தி பனுவல்களில் உள்ள தவறான சொற்களைக் கண்டுபிடிக்க உதவும் மற்றும் தவறான சொற்களுக்குச் சரியான மாற்றுச் சொற்களைத் தரும். முதல் செயல்பாடு “பிழை கண்டுபிடித்தல் (error detection)” எனப்படும்; இரண்டாவது செயல்பாடு “பிழைதிருத்தல் அல்லது பரிந்துரை உருவாக்குதல்” (error correction or suggestion generation) எனப்படும்.

4.1.3.1 எழுத்துப்பிழைதிருத்தலின் பலவேறு அணுகுமுறைகள்

விஜய் பின்வருவனவற்றை எழுத்துப்பிழை திருத்தலின் பல்வேறு அணுகுமுறைகளாகப் பட்டியலிடுகின்றார்.

- ஒலியல் அடிப்படையில் ஒன்றுபோல் இருக்கும் எழுத்துகளின் பிழைதிருத்தம்
- பிழையைக் கண்டுபிடித்தல் மற்றும் இருப்பிடம் கண்டுபிடித்தல்
- ஓசை ஊடக மாதிரி (noise channel model)
- ஒட்டுநிலை மொழிகளின் எழுத்துப்பிழை திருத்தம் (ஒஃபல்ஸர், கே 1994)
- எழுத்துப்பிழை திருத்த அணுகுமுறைக்குச் சாளர அடிப்படை அணுகுமுறை (Window based approach)
- எழுத்துப்பிழை திருத்தத்தில் நிகழ்தகழ்வு மதிப்பெண்

4.1.3.2 இந்திய மொழிகளுக்கு இருக்கின்ற எழுத்துப் பிழை திருத்திகள்

பின்வரும் அட்டவணையில் இந்திய மொழிகளுக்கு இருக்கின்ற எழுத்துப் பிழை திருத்திகளை விஜய் பட்டியலிடுகின்றார்.

மொழிகள்	உருவாக்கப்பட்ட இடம்	பயன்படுத்தப்பட்ட அணுகுமுறை
அசாம் மொழி	RICILTS, கௌகாத்தி	அகராதி அடிப்படை அணுகுமுறை
வங்காள மொழி	இந்தியப் புள்ளியியல் நிறுவனம், கொல்கொத்தா (பி.பி. சௌத்ரி)	
தெலுங்கு கன்னட மொழிகள்	தெலுங்கு மூலவள மையம், ஹைதராபாத் பல்கலைக் கழகம்	உருபனியல் ஆய்வி அடிப்படை அணுகுமுறை
மலையாளம்	RICILTS, திருவனந்தபுரம்	விதி மற்றும் அகராதி அடிப்படை அணுகுமுறை
மராத்தி	RICILTS, ஐஐடி பாம்பே	அகராதி மற்றும் பைகிராம்

		மற்றும் டிரைகிராம் அணுகுமுறை
தமிழ்	தனபாலன், அண்ணா பல்கலைக் கழகம், எயுகேபிசி, சென்னைப் பல்கலைக் கழகம்	உருபனியல் பகுப்பாய்வியைப் பயன்படுத்தி தரவுத்தொகுதிஅடிப்படை + உருபனியல் பகுப்பாய்வி விதிகள்

4.1.3.3 பிழை கண்டுபிடித்தல்

இதை இருவழிகளில் செய்ய இயலும்

1. சொல்லை அகராதியில் ஒப்பிட்டுப் பார்த்தல் (அகராதி அணுகுமுறை)
2. சொல்லை உருபனியல் ஆய்வியால் ஆய்ந்து பிழை கண்டுபிடித்தல் (மொழியியல் அணுகுமுறை)

இரண்டு அணுகுமுறைகள் உள்ளன :

1. தரவுத்தொகுதி அடிப்படையிலான அணுகுமுறை (முற்றுநிலைத் தானியங்கியைப் பயன்படுத்தி)
2. மொழியியல் அடிப்படையிலான அணுகுமுறை (உருபனியல் ஆய்வியைப் பயன்படுத்தி)

4.1.3.3.1 தரவுத்தொகுதி அடிப்படையிலான அணுகுமுறை

தரவுத்தொகுதி முற்றுநிலைத் தானியங்கியால் (Finite State Automata) உருப்படுத்தம் செய்யப்படும். இது விரைவக மதிப்பீடு செய்யும். ஒரு கச்சிதமான உருப்படுத்தத்தைச் செய்யும். சொல்லின் ஒவ்வொரு எழுத்தும் முற்றுநிலைத் தானியியங்கியில் நிலைமாற்றக் குறியீடாக (transition symbol) இருக்கும்.

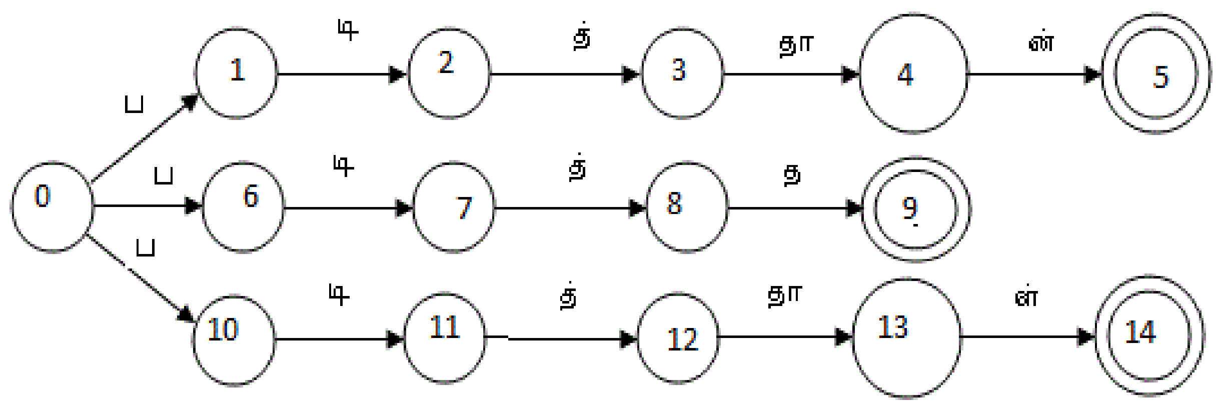
முற்றுநிலைத் தானியங்கி

முற்றுநிலைத் தானியங்கி பனுவல் கோர்வையில் எளிதான தொடரியல் அமைப்புகளை அல்லது ஒழுங்கு அமைப்புகளைப் புரிந்துகொள்வதற்காகப் பயன்படுத்தப்படும் அருவ உபாயம் ஆகும். ஒரு தனியிங்கி நிலை படம் (state diagram)

என்று அழைக்கப்படுகின்ற வழிகாட்டப்பட்ட வரைபடத்தால் (directed graph) வழக்கமாக விவரிக்கப்படும். முற்றுநிலைத் தானியங்கி ஒரு கோர்வை ஆய்வு உபாயம் (string processing device) ஆகும்; அது ஒரு பனுவல் கோர்வையை உள்ளீடாக ஏற்கும் மற்றும் கோர்வையை ஏற்கவோ மறுக்கவோ செய்யும். கணித அடிப்படையில் இதை ஒரு குழுமத்திற்கு {ஏல், மறு} ஒரு குழுமக் கோர்வைகளைப் பொருத்தும் ஒரு செயல்பாடாகக் கருதலாம்.

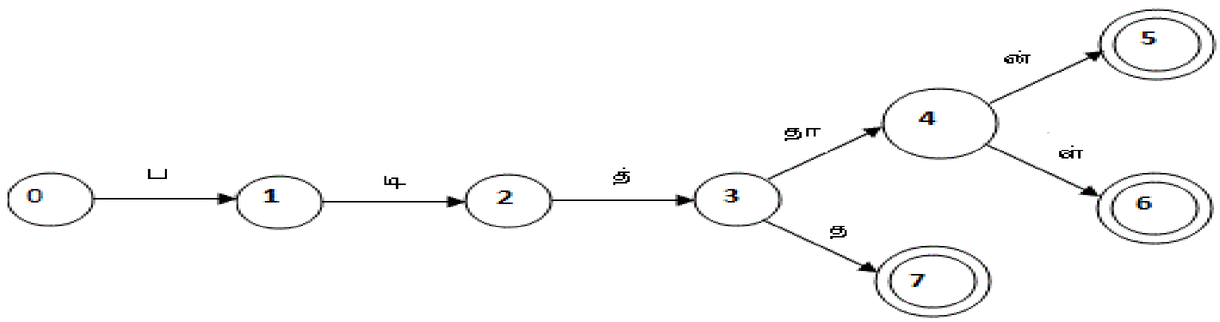
நிலைகள் வட்டங்களால் உருப்படுத்தம் செய்யப்படும். உள்ளீடாக ஒரு குறியீடு உபாயத்தால் பெறப்பட்டால் நேர்கின்ற நிலைமாற்றங்களை (transitions) அம்புக்குறிகள் உருப்படுத்தம் செய்யும். தொடக்கத்தில் தீர்மானிக்கப்படாத (non-diterministic) முற்றுநிலைத் தானியங்கி உருவாக்கப்பட்டது; இது பின்னர் முற்றுநிலைத் தானியங்கி கருவித்தொகுதியைப் (tool kit) பயன்படுத்தித் தீர்மானிக்கப்பட்ட (diterministic) மற்றும் குறுக்கப்பட்ட (minimized) முற்றுநிலைத் தானியங்கியாக ஆக்கப்பட்டது.

தீர்மானிக்கப்படாத முற்றுநிலைத்தானியங்கி



ஒரே குறியீடுகள் பல நிலைகளுக்குத் தூண்ட இயலும். பல வழிகள் சாத்தியமாகும். கணினிச் செயல்பாடுகளுக்குப் பொருந்தாது.

தீர்மானிக்கப்பட்ட முற்றுநிலைத் தானியங்கி



ஒரு குறியீடு ஒரு நிலையிலிருந்து மற்றொரு நிலைக்கு மட்டுமே தூண்ட இயலும். பல வழிகள் இல்லை. இதன் அனுகூலம் முனைப்பான உருப்படுத்தமும் விரைவான செயற்பாங்கும் ஆகும்.

நிலை அட்டவணை உருப்படுத்தம்

ஆரம்பநிலை	**அடுத்த நிலை**	**நிலைமாற்றக் குறியீடு**
0	1	ப
1	2	டி
2	3	த்
3	4	தா
3	7	த
7	இறுதிநிலை	
4	5	ன்
4	6	ள்
5	இறுதிநிலை	
6	இறுதிநிலை	

4.1.3.3.2 மொழியியல் அடிப்படையிலான அணுகுமுறை

உருபனியல் பகுப்பாய்வி

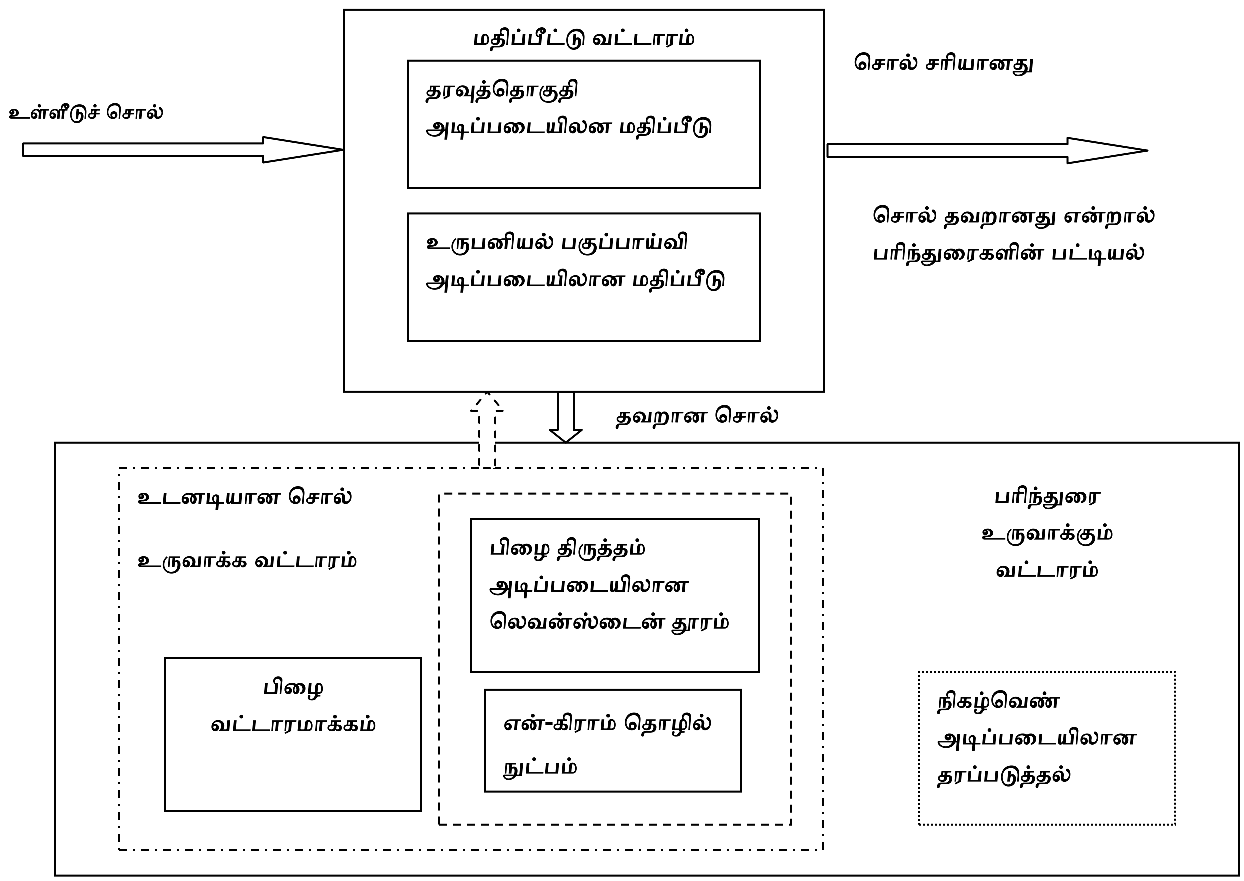
உருபனியல் பகுப்பாய்வு என்பது சொல்லை அதில் உள்ளடங்கும் உருபன்களாகப் பிரிதல் ஆகும். இச்செயலில் உருபன்களின் அடையாளங்கள் தேவையில்லை. சொல்லுக்கு ஏற்கத்தக்க பகுப்பாய்வு இருந்தால் அது ஏற்கத்தக்கச் சொல்லாகக் கருதப்படும்.

படித்தான் = படி + த்த் + ஆன்

பின்னொட்டு நீக்கி (suffix stripper) இச்செயன்மைக்குப் போதுமானதா?

4.1.3.4 பிழை திருத்தியின் கட்டமைப்பு

பிழைதிருத்தியின் கட்டமைப்பு கீழே வரைக்கப்பட்டு விளக்கப்பட்டுள்ளது.



4.1.3.4.1 பிழை திருத்தல்

பிழை திருத்துவதற்குப் சவுண்டெக்ஸ் கொள்கை (soundex principle) போன்ற பல அணுகுமுறைகள் பின்பற்றப் படுகின்றன. சொல் தவறு என்று கண்டுபிடிக்கப்பட்டால் பின்வருவனவற்றைப் பயன்படுத்தி பரித்துரைகள் உருவாக்கப்படுகின்றன.

- பிழையை வட்டார எல்லைக்கு உட்படுத்தல்
- திருத்த தூரத்தையும் என்கிராமையும் பயன்படுத்தி இடைப்பட்ட சொற்களை உருவாக்குதல்
- அதை மதிப்பிடுதல்
- தரப்படுத்துதல்

பிழை திருத்தம்

இது பின்வரும் தொழில் நுட்பங்களை உள்ளடக்கும்

1. லெவன்ஸ்டைன் தூரம்
2. பிழையை எல்லைக்குட்படுத்தல்
3. என்-கிராம் தொழில் நுட்பம்
4. நிகழ்வெண் அடிப்படையிலான தரப்படுத்தும் படிமுறை

4.1.3.4.1.1 லெவன்ஸ்டைன் தூரம்

லெவன்ஸ்டைன் தூரம் (Levenstein distance) என்பது வலாடிமிர் லெவன்ஸ்டைன் என்ற ரஷ்ய அறிவியல் அறிஞரின் நினைவாகப் பெயரிடப்பட்டுள்ளது. பதின்ம அளவுமுறையும் சிலவேளைகளில் திருத்து தூரம் (edit distance) என்று அழைக்கப்படும். லெவன்ஸ்டைன் தூரம் (LD) என்பது இரு கோர்வைகளுக்கு இடையிலுள்ள ஒற்றுமை ஆகும்; இதில் நாம் மூலக் கோர்வையை s என்றும் இலக்குக் கோர்வையை t என்றும் குறிப்பிடுகின்றோம்.

t ஆக மாற்றத் தேவைப்படும் நீக்கல் (deletion), செருகல் (insertion) அல்லது பதிலீடுசெய்தல் (substitution) என்ற நடவடிக்கையின் எண்ணிக்கையின் வடிவில் தூரம் வெளிப்படுத்தப்படுகின்றது. எடுத்துக்காட்டாக, s என்பது “test” மற்றும் t என்பது

“test” என்றால் LD (s,t) = 0, ஏனென்றால் எந்த மாற்றமும் தேவையில்லை. இரண்டு கோர்வைகளும் ஒன்றாக இருக்கின்றன. s என்பது “test” மற்றும் t என்பது “tent” என்றால் LD (s,t) = 1, ஏனென்றால் s-ஐ t-ஆக மாற்ற ஒரு பதிலீடு முறை (“s”-ஐ “n”–ஆக மாற்ற) போதுமானது.

மூலச் சொல்: படித்த

இலக்குச்சொல்: படித்தான்

லெவன்ஸ்டைன் தூரம் 2

மூலச் சொல்: படித்தான்

இலக்குச் சொல்: படித்தாள்

லெவன்ஸ்டைன் தூரம் 1

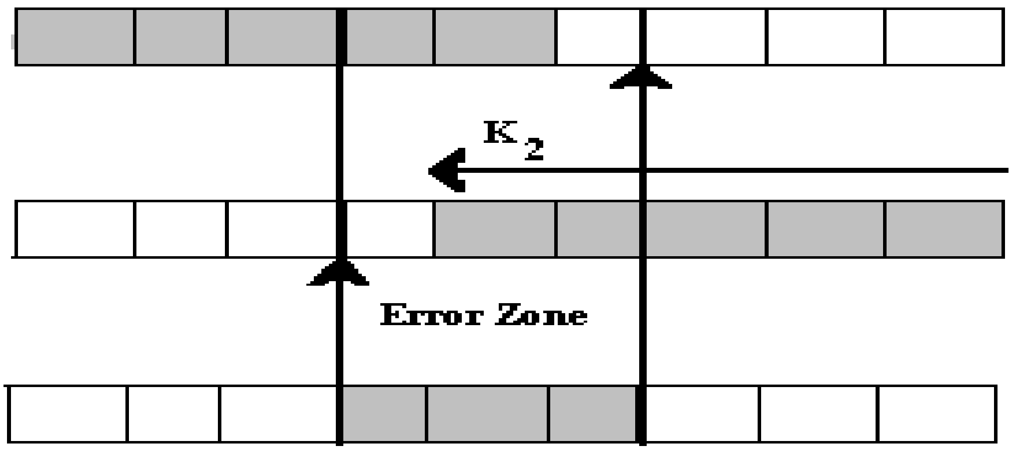
4.1.3.4.1 2 பிழை எல்லைப்படுத்தல்

பிழை செய்யப்பட்ட சொல்லின் பிழை பரப்பைக் (error zone) கண்டுபிடிக்கும் நுட்பம் பிழை எல்லைப்படுத்தல் எனப்படும். கட்டுப்பாட்டிற்குட்பட்ட செயற்பாங்கை நிறைவேற்றவும் எதிர்பார்க்கப்பட்ட மாற்றுப் பரிந்துரைகளை உருவாக்கவும் பிழை எல்லைப்படுத்தம் தேவையாகும்.

நாம் நமது செயல்பாட்டில் பின்வரும் ஊகங்களைச் செய்கின்றோம்:

1. சொல்லில் செருகல், நீக்கல், பதிலீடு செய்தல், இடம் மாறுதல் என்பனவற்றுள் ஒரு தவறு தான் இருக்கக்கூடும்.
2. சரியான சொல் இரு அகராதிகளிலும் (மரபு மற்றும் திருப்பப்பட்ட சொல்) கோப்புகளிலும் இருக்கும்.

D_c என்ற மரபு அகராதியும் R_c என்ற மாற்றப்பட்ட அகராதியும் பயன்படுத்தப்படுகின்றன. $W = x_1x_2..x_k$ என்ற சொல்லின் மாற்றப்பட்ட சொல் $W_r = x_kx_{k-1}…x_2x_1$ ஆகும், இதில் x_1 என்பது ஏற்கத்தக்க தமிழ் எழுத்தாகும். R_c என்ற மாற்றப்பட்ட அகராதியில் மரபு அகராதி D_c-யின் எல்லாச் சொற்களின் திருப்பப்பட்ட பதிப்புருவும் வைத்திருக்கப்படும்.



மரபு மற்றும் மாற்ற அகராதி அடிப்படையில் பிழை எல்லைப்படுதல்

இது மூன்று விதமாகச் செய்யப்படுகின்றது.

(அ) மரபு அகராதி மூலம் பிழை எல்லைப்படுத்தல்

(ஆ) திருப்பு அகராதி மூலம் பிழை எல்லைப்படுத்தம்

(இ) இரு அகராதியும் பயன்படுத்தப்படுகையில் பிழை எல்லைப்படுத்தம்

சரியான சொல்: குதிரைகள்

தவறான உள்ளீட்டுச் சொல்: குதரகள்

திருப்புப் பகுப்பு

கு	த	ரை	க	ள்
முன்னோக்கிய பகுப்பு				
கு	த	ரை	க	ள்
பிழை பரப்பு				
கு	த	ரை	க	ள்

4.1.3.4.1.3 தவறு பரப்பில் திருத்தம்

இது என்-கிராம் நுட்பத்தைப் பயன்படுத்திச் செய்யப்படுகின்றது. என்-கிராம் இலக்கணம் என்-ஆவது நிரல் மார்க்கொவ் மொழி மாதிரியின் உருப்படுத்தம் ஆகும்; இதில் ஒரு குறியீட்டின் நேர்வு சாத்தியம் பிற குறியீட்டின் N -1-ஆவதன் முந்தைய நேர்வால் கட்டுப்படுத்தப்பட்டுள்ளது. ஒரு W என்ற தரப்பட்ட சொல் வரிசையின் காரணகாரிய சாத்தியத்துடன் P(W) புரிந்துகொள்வானைத் தருவதற்கு என்-கிராம் மாதிரிகள் பெரிய சொற்றொகை பேச்சு புரிந்துகொள்ளல் ஒழுங்கமைப்புகளில் பயன்படுத்தப்படுகின்றது. என்–கிராம் மொழி மாதிரிகள் எதிர்பார்க்கப்பட்ட உள்ளீடாக மொழியின் ஒரே பண்புகளைப் பங்கிட்டுக்கொள்ளும் பெரிய பயிற்சிப் பனுவல்களிலிருந்து வழக்கமாக ஆக்கப்படுகின்றது. இந்தத் தகவல் பேசுபவர்களின் பேசு உருவாக்கிகளின் பண்புக்கூறுகளை மாதிரிசெய்யும் ஒலியியக்கவியல் மாதிரி P(W|O) என்பதை குறைநிறவாக்கம் செய்யும். இக்கூறுகள் இரண்டும் இணைந்து $W' = armax_w$ P(W|O) என்ற பெரும்பான்மை சாத்தியமான உள்ளீட்டு வரிசையை கணிக்க ஒரு ஒழுங்குமுறையை அனுமதிக்கும்; இதில் $W' = armax_w$ P(W|O) P (W) ஆக O என்பது உள்ளீட்டு சமிக்கையாகும். நாம் என்-கிராமை சொல்நிலையில் அல்லாமல் எழுத்துநிலையில் பயன்படுத்துகின்றோம்; நாம் N= 1,2,3 என்று எடுத்துக்கொள்கின்றோம். பைகிராம் நுட்பத்தில் முதலில் வரும் எழுத்துக்கு எல்லாச் சாத்தியமான அடுத்த எழுத்தைக் கண்டுபிடிக்க முயலுகின்றோம். டிரைகிராம் நுட்பத்தில் ஒரு குறிப்பிட்ட முதல் மற்றும் மூன்றாவது எழுத்துக்குச் சாத்தியமான எல்லா மைய எழுத்தைக் கண்டுபிடிக்க முயலுகின்றோம்.

4.1.3.4.1 4 ஆக்கப்பட்ட சொற்களுக்களின் மதிப்பீடு

பிழை எல்லைப்படுத்தல் நுட்பங்களைப் பயன்படுத்தி பிழை பரப்புகள் அடையாளம் காணப்படும். என் கிராம் நுட்பத்தைப் பயன்படுத்தி இடைநிலைச் சொற்கள் உருவாக்கப்படும். இந்த இடைநிலைச் சொற்கள் பின்வருவனவற்றைப் பயன்படுத்தி மதிப்பீடு செய்யப்படும்:

- முற்றுநிலைத் தானியங்கி அடிப்படையிலான தரவுத்தொகுதி அணுகுமுறை
- உருபனியல் பகுப்பாய்வி

4.1.3.4.1.5 உருவாக்கப்பட்ட சொற்களின் முன்னிலைப்படுத்தல்

உருவாக்கப்பட்ட சொற்கள் தரப்படுத்தப்படவேண்டும். தரப்படுத்தம் பின்வரும் மூன்று அடிப்படையில் அமையும்:

- நிகழ்வெண் அடிப்படையில்
- பயன்பாட்டாளர் அடிப்படையில்
- பொருண்மைக்களம் அடிப்படையில்

4.1.3.4.1 .6 பிழை திருத்தியின் பயன்பாடு

பனுவல் செம்மையாக்கிகளில் பயன்படுத்தப்படுகின்றது. தேடல் இயந்திரங்களிலும் போர்டல்களிலும் பயன்படுத்தப்படுகின்றது.

4.1.4 தெய்வசுந்தரத்தின் மென் தமிழ் சொல்லாளர்

தெய்வசுந்தரம் தமது மொழியியல் அறிவையும் கணிப்பொறி மொழியியல் அறிவையும் பயன்படுத்தி பல்லாண்டு முயற்சிகளுக்குப் பின் மிகச் சிறப்பான ஒரு தமிழ் சொல்லாய்வியை உருவாக்கியுள்ளார். இது சொற்பிழை திருத்தி, சந்திப்பிழை திருத்தி, மொழிமாற்றி என்ற பகுதிகளைக் கொண்டது. இது இயல்மொழி ஆய்வு, கணினி மொழியியல் ஆய்வு, மொழித்தொழில் நுட்பம் ஆகியவற்றின் அடிப்படையில் உருவாக்கப்பட்ட ஒரு மென்மம் ஆகும். இதில் தமிழை உள்ளீடு செய்ய பலவகையான விசைப்பலகைகள் உள்ளன. ஒருங்குறி (Unicode) எழுத்துருக்களைப் பயன்படுத்தும் வசதி உள்ளது. எந்த வகைக் குறியத்தையும் வடிவமைப்பு மாறாமல் எந்தக் குறித்துக்கும் மாற்றலாம். அகராதிச் சொற்கள் மட்டுமல்லாமல் விகுதிகள் ஏற்ற சொற்களையும் ஆங்கிலச் சொற்களையும் திருத்தித்தரும். ஒற்று (சந்தி) இடாத இடங்களில் இடும். தேவையில்லாத இடங்களில் ஒற்றினை நீக்கியும் தரும். ஆங்கிலச் சொற்களை ஒலிபெயர்த்து எழுதினாலும் அவற்றுக்கு நிகரான தமிழ்ச் சொற்களாக மாறித்தரும். ஆறுவகையான அகரதிகள் உள்ளன: 45 ஆயிரம் தற்காலத் தமிழ்ச் சொற்களைக் கொண்ட தமிழ்-ஆங்கிலம், ஆங்கிலம்-தமிழகராதி, இணைச் சொல் அகராதி (இது ஒரு சொல்லுக்கு

இணையான பல சொற்களை வழங்கும்), எதிர்சொல் அகராதி (இது ஒரு சொல்லுக்குரிய எதிர்ச்சொற்களை வழங்கும்), மயங்கொலிச்சொல் அகராதி (தமிழில் ஏற்படும் *ல*, *ள*, *ழ* மயக்கங்கள், *ந*, *ண*, *ன* மயக்கங்கள், *ர*, *ற* மயக்கங்கள் ஆகியவற்றைத் தவிர்க்கப் பயன்படும் பொருள் வேறுபாட்டுக் குறிப்புகள் அடங்கியது), தமிழ்நாட்டு அரசு ஆட்சிச்சொல் அகராதி. மேலும் அகரவரிசைப்படுத்தி, சொல்லடைவு, எண்-எழுத்து மாற்றி (பயன்பாட்டில் இருக்கும் எத்தகைய எண்களையும் தமிப் புணர்ச்சி விதிப்படி தமிழ்சொற்களாக மாற்றித்தரும், கோப்புமாற்றி போன்ற பலவித வசதிகளைக் கொண்டு இயங்குவது.

4.2. ஒரு மாதிரிப் பிழை திருத்தி

இங்கு ஒரு மாதிரி எழுத்து மற்றும் இலக்கணத் திருத்தி உருவாக்குவது குறித்த தொழில் நுட்பம் கூறப்பட்டுள்ளது. இது தரவுத்தொகுதி அடிப்படையிலான பிழைதிருத்தி. இங்கு தரப்பட்டுள்ள தரவுத்தொகுதி அடிப்படையில் இது இயங்கும்.

4.2.1 பிழைதிருத்தியின் வழிமுறை அமைப்பு

சொல் தரப்படுகையில் நாம் அச்சொல்லின் சாத்தியமான எழுத்துப் பிழை திருத்தத்தைத் தேர்ந்தெடுக்க முயலுவோம். (அந்தச் சரியான சொல் மூலச்சொல்லாகவே இருக்கலாம்.) இதை உறுதியாக அறிவதற்கு வழியில்லை; எனவே நாம் சாத்தியங்களைப் பரிந்துரைக்கின்றோம். எடுத்துக்காட்டாக சாத்தியமான எல்லா திருத்தங்களிலிருந்தும் சரியான c என்ற திருத்தத்தைக் கண்டுபிடிப்பதற்கு முயலுவோம்; இது மூலச்சொல் தரப்படுகையில் c-யின் சாத்தியங்களை மேம்படுத்தும்:

argmaxc P(c|w)

By Bayes' Theorem this is equivalent to:

argmaxc P(w|c) P(c) / P(w)

Since P(w) is the same for every possible c, we can ignore it, giving:

argmaxc P(w|c) P(c)

இந்த வெளிப்படுத்தத்தின் மூன்று பகுதிகள் உள்ளன. வலமிருந்து இடமாக நமக்குப் பின்வருவது இருக்கின்றது:

1.P (c), பரிந்துரைக்கப்பட்ட திருத்தம் c தானாகவே நிலைத்துநிற்கும் என்ற சாத்தியம். இது மொழி மாதிரி என்று அழைக்கப்படும்: “தமிழ்ப் பனுவலில் c தோன்றுவது எந்த அளவுக்குச் சாத்தியம்?” என்ற கேள்விக்கு விடை அளிப்பதாக எண்ணிக்கொள்ளவும்; P("zxzxzxzyyy") என்பது பூஜியத்திற்கு அண்மைப்படும்.

2. P(w|c), ஒருவர் c-ஐ குறிப்பிட எண்ணும் போது w என்பது தட்டச்சு செய்யப்படும் சாத்தியம். இது பிழை மாதிரியாகும்: ”ஒருவர் c-ஐ குறிப்பிட விரும்புகையில் தவறாக w-ஐத் தட்டச்சு செய்யவது எந்த அளவுக்குச் சாத்தியமானது?” என்ற கேள்விக்கு விடையளிப்பதாக எண்ணிக்கொள்ளவும்.

3. கட்டுப்பாடு உபாயம், argmaxc,

3.argmaxc, c-யின் எல்லாச் சாத்தியமான மதிப்புகளைப் பட்டியலிடவும் மற்றும் மிகச் சரியான ஒருங்கிணைந்த சாத்திய மதிப்பெண்ணை தெரிந்தெடுக்கவும் கூறும் கட்டுப்பாட்டு உபாயம்.

வெளிப்படையான கேள்வி பின்வருவதாகும்: நாம் P(c/w) என்ற எளிய வெளிப்படுத்தத்தை எடுத்துக்கொண்டு ஒரு மாதிரிக்குப் பதிலாக இரு மாதிரிகளை உட்படுத்தும் கூடுதல் கலவைத் தன்மையான வெளிப்பாட்டால் இடம்பெயர்க்கின்றோம்? இதன் விடை (P(c/w) ஏற்கனவே இரு காரணிகளை ஒருங்கிணைக்கின்றது மற்றும் இரண்டையும் பிரிப்பதும் அவற்றை வெளிப்படையாகக் கையாளுவதும் எளிது. w= “பட்தம்” என்ற தவறான சொல்லைக் கருத்தில் கொள்ளவும்; இதற்கு குறைந்தது இரண்டு திருத்தங்கள் இருக்கலாம்: c = படம், c = பட்டம். இவற்றில் எதற்குக் கூடுதல் P(c/w) மதிப்பு இருக்கின்றது? ”த” என்பதை “ட” என மாற்றினால் போதுமானது என்ற சிறிய மாற்றம் என்ற அளவில் “பட்டம்” சரியானதாகத் தோன்றுகின்றது. மாறாக, “படம்” என்பது மிகப் பொதுவாகப் பயன்படுத்தப்படுகின்ற சொல் என்ற நிலையிலும் விசைப் அண்மை என்ற நிலையிலும் படம் என்பதைத் தவறாக தட்டச்சு செய்ய வாய்ப்பு இருக்கின்றது. முக்கியமாத விசயம் நாம் c என்பதன் சாத்தியத்தையும் c என்பது w என மாறுவதன் சாத்தியத்தையும் கருத்தில் கொள்ளவேண்டும்; எனவே இவ்விரு காரணிகளைப் முறையாகப் பிரிப்பது தெளிவானது. இபோது நாம் எவ்வாறு வழிமுறை வரைவு வேலை செய்கின்றது என்பதைக் காட்டுவதற்குத் தயாராய் உள்ளோம்.

நாம் ஒரு பெரிய பனுவல் கோப்பைப் (big.txt) படிப்போம்; இது ஒரு மில்லியன் சொற்களைக் கொண்டிருக்கும். இது பல வித தரவுகள் சேர்ந்த தொகுதியாகும்.

பின்னர் நாம் கோப்பிலிருந்து தனிச் சொற்களைப் பிரித்தெடுப்போம் (இதன் படி ‘படம்’, ‘பட்டம்’ என்ற சொற்கள் ஒன்றாகும் மற்றும் சொற்கள் எழுத்துக்களின் கோர்வையாகும்.) அடுத்தபடியாக சாத்திய மாதிரியைப் பயிற்சிசெய்யவும்; இது செயல்பாட்டுப் பயிற்சையைப் பயன்படுத்தி எவ்வளவு தடவை ஒவ்வொரு சொல்லும் வருகின்றது என்பதைக் கூறுவதாகும். பின்வருதைப் பார்க்கவும்:

```
def words(text): return re.findall('[a-z]+', text.lower())
def train(features):
    model = collections.defaultdict(lambda: 1)
    for f in features:
        model[f] += 1
    return model
NWORDS = train(words(file('big.txt').read()))
```

இந்த இடத்தில் NWORDS[w] என்பது w எத்தனை தடவை காணப்படுகின்றது என்ற கணக்கைக் கொண்டிருக்கின்றது. ஒரு தொகுப்பு இருக்கின்றது: புதிய சொற்கள். தமிழில் நல்ல சொல் பயிற்சித் தரவில் இல்லாவிடில் என்ன நடக்கும் என்பது கேள்வி. இச்சிக்கலுக்குப் பல தரமான அணுகுமுறைகள் உள்ளன; நாம் எளிதான ஒன்றை எடுக்கின்றோம்; இது புதிய சொற்களை நாம் அவற்றை ஒருதடவைப் பார்த்திருப்தாகக் கையாளும். இந்த பொதுவான செயல்முறை சீராக்கம் (smoothing) எனப்படும்; ஏனென்றால் நாம் பூஜியமாக இருக்க இயலும் சாத்திய வினியோகத்தின் பகுதிகளை மிகக் குறைவான சாத்திய எண்ணிக்கைக்கு மாறும் படிக்கு சீர்மை செய்வோம். இது வகுப்புச் சேகரிப்புகள், defaultdict வழியாக எய்தப்படுத்துகின்றது; இது பைத்தான் dict (பிற மொழிகள் இதை hash table என்று அழைக்கும்) போன்று சீரானது; நாம் எந்த விசையின் வழுநிலை மதிப்பையும் குறிப்பிட இயலும் என்பது விதிவிலக்காகும்; இங்கு நாம் 1-ஐப் பயன்படுத்துகின்றோம்.

இப்போது நாம் w என்ற சொல்லின் சாத்தியமான திருத்தங்கள் c-ஐப் பட்டியலிடும் சிக்கலைப் பார்ப்போம். இரு சொற்களுக்கு இடையிலுள்ள திருத்து தூரத்தை (edit

distance) பற்றி பேசுவது பொதுவானதாகும்: ஒன்றை மற்றொன்றாக மாற்றுவதற்கு அது எடுக்கும் திருத்தங்களின் எண்ணிக்கை. நிக்குவது (ஒரு எழுத்தை நீக்குவது), இடம்மாற்றுவது (அண்மையில் வரும் எழுத்துக்களை இடம் மாற்றுவது), மாற்றுவது (ஒரு எழுத்தை மற்றொரு எழுத்தாக மாற்றுவது) என திருத்தம் அமையலாம். கீழே வருவது w என்பதிலிருந்து ஒரு திருத்த தூரத்தில் உள்ள எல்லா c சொற்களின் குழுமத்தை திருப்பி அனுப்பும் செயல்பாடாகும்.

```
def edits1(word):
  splits     = [(word[:i], word[i:]) for i in range(len(word) + 1)]
  deletes    = [a + b[1:] for a, b in splits if b]
  transposes = [a + b[1] + b[0] + b[2:] for a, b in splits if len(b)>1]
  replaces   = [a + c + b[1:] for a, b in splits for c in alphabet if b]
  inserts    = [a + c + b     for a, b in splits for c in alphabet]
  return set(deletes + transposes + replaces + inserts)
```

இது ஒரு பெரிய குழுமம் ஆகும். n நீளம் உள்ள ஒரு சொல்லுக்கு n நீக்கங்கள், n-1 இடமாற்றங்கள், 26n மாற்றங்கள், 26(n+1) செருகல்கள் இருக்கும்; மொத்தம் 54n+25 ஆகும் (இவற்றில் சில இரட்டிப்புகளாகும். எடுத்துக்காட்டாக, len(edits1 ('something'))--, அதாவது edits1('something') என்பதன் தனிமங்களின் எண்ணிக்கை 494 ஆகும்.

பிழை திருத்தம் பற்றிய இலக்கியங்கள் எழுத்துப் பிழைகளின் 80%-இலிருந்து 90% வரை இலக்கிலிருந்து 1 திருத்து தூரத்தில் இருக்கின்றது என்று கூறுகின்றன. நாம் 2 திருத்து தூரத்தைக் கருத்தில் கொள்ளவது தேவையாகும். அது எளிதாகும்: எல்லா edits1-இன் விளைவுகளுக்கும் edits1-ஐப் பிரயோகிக்கவும்.

```
def edits2(word):
return set(e2 for e1 in edits1(word) for e2 in edits1(e1))
```

இது எழுதுவதற்கு எளிதாகும், ஆனால் நாம் சில ஆபத்தான சிக்கல்களில் சிக்கிக்கொள்வதற்குத் தொடங்குகின்றோம்: len(edits2('something')) is 114,324. இருப்பினும் நமக்கு நல்ல செயலெல்லை கிடைக்கின்றது; 270 பரிசோதனை நேர்வுகளில் 3 தான் 2-ஐவிட கூடுதல் திருத்து தூரத்தைக் கொண்டுள்ளன. அதாவது edits2 நேர்வுகளின் 98%-ஐ உட்படுத்துகின்றது; அதாவது அது போதுமானதாகும். நாம் திருத்து தூரம் 2-ஐத் தாண்டி செல்லமாட்டோம் என்பதால் நாம் சிறிய உகப்பாக்கத்தைச் செய்ய இயலும்: சொற்கள்

என்று உண்மையில் கூறப்படும் கருதத்தக்கவைகளை (candidates) மட்டும் வைத்திருக்கவும்; நாம் இன்னும் எல்லாச் சாத்தியங்களையும் கருத்தில் கொள்ளவேண்டும்; ஆனால் நாம் அவற்றின் பெரிய குழுமத்தைக் கட்டத் தேவை இல்லை. known_edits2 என்ற செயல்பாடு பின்வருவதைச் செய்யும்:

```
def known_edits2(word):
return set(e2 for e1 in edits1(word) for e2 in edits1(e1) if e2 in NWORDS)
```

எடுத்துக்காட்டாக இப்போது known_edits2('something') 4 சொற்களின் குழுமமாகும்: edits2-ஆல் உருவாக்கப்பட்ட 114,324 சொற்களின் குழுமத்திற்குப் பதிலாக {'smoothing', 'seething', 'something', 'soothing'}ஆகும். இது காரியங்களை கிட்டத்தட்ட 10% வரை விரைவுபடுத்தும்.

இப்பொழுது விடுபட்ட ஒரே ஒரு பகுதி பிழை மாதிரி P(w/c) ஆகும். இங்கு நாம் சிக்கலில் அகப்படுவோம். திருத்து தூரம் 1-இன் எல்லா தெரிந்த சொற்களும் திருத்து தூரம் 2-இன் தெரிந்த சொற்களை விட முடிவற்ற கூடுதல் சாத்தியமாகும். மற்றும் திருத்து தூரம் 0-இன் தெரிந்த சொற்களை விட குறைத்த சாத்தியமாகும். 'தெரிந்த சொல்' என்பது மொழி மாதிரி பயிற்சி தரவில் நாம் கண்டிருப்பவை -- அகராதியில் உள்ள ஒரு சொல். நாம் இந்த நடவடிக்கையைப் பின்வருமாரு செயல்படுத்தலாம்:

```
def known(words): return set(w for w in words if w in NWORDS)
def correct(word):
candidates = known([word]) or known(edits1(word)) or known_edits2(word) or [word]
return max(candidates, key=NWORDS.get)
```

திருத்து என்ற செயல்பாடு (function correct) மூலச் சொல்லுக்கு குறைந்த திருத்து தூரம் இருக்கிற கருதத்தக்கச் சொற்களின் (candidate words) குழுமத்தை தெரிந்தெடுக்கின்றது; அந்த குழுமத்தில் சில தெரிந்த சொற்கள் (known words) இருப்பதுவரை இது நடைபெறும். இது கருத்தில் கொள்ளத்தக்க குழுமத்தை அடையாளம் அடையாளம் கண்டுகொண்டால் அது NWORDS மாதிரியால் மதிப்பிடப்பட்ட மிகக்கூடுதல் P(c) மதிப்புள்ள தனிமத்தைத் தெர்ந்தெடுக்கும்.

மேற்கூறியவாறு தமிழ் எழுத்துப்பிழை மற்றும் இலக்கணப் பிழை திருத்தி உருவாக்கப்பட்டு தட்டச்சு செய்யபட்ட பனுவல்களை திருத்தப் பயன்படுத்தப்படும்.

இத்தகைய முயற்சி இங்கு ஒரு மாதிரிக்காக மேற்கொள்ளப்பட்டுள்ளது. முன்னர் கூறியது போன்று இவ்வாய்வு ஒரு பரிசோதனை என்ற நிலையிலேயெ அமைகின்றது.

4.3. முடிவுரை

தமிழில் வெளிவந்த எழுத்துப் பிழை மற்றும் இலக்கணப் பிழை திருத்திகளின் பல்வேறுபட்ட பரிமாணங்கள் குறித்து இங்கு விளக்கப்பட்டது. சொல்லாய்விகளும் சொல்லாளர்களும் பிழை திருத்திகளை உள்ளடக்கி வெளிவரத் தொடங்கின. தொழில் நுட்பவியலார்களும் மொழியியலாளர்களும் இதற்கென்று முக்கியத்துவம் கொடுத்து பல நிலைகளில் எழுத்து மற்றும் இலக்கணத் திருத்திகளை உருவாக்க முயன்றனர். மொழியியல் விதிகளை உள்ளடக்கி திருத்திகளை உருவாக்கச் சிலர் முயன்றனர். இதில் குறிப்பிடத் தகுந்தவர் தெய்வசுந்தரம் அவர்கள். இந்த விதிகளை அறியாத அல்லது விதிகளைப் பற்றி கவனம் செலுத்தாது உடனடியான எழுத்துப் பிழை திருத்தியும் இலக்கணப் பிழை திருத்தியும் உருவாக்க முயன்றவர்களில் பலர் தரவுத்தொகுதியைப் பயன்படுத்தி இத்தகைய பிழை திருத்திகளை உருவாக்க முயன்றனர். எல்லா முயற்சிகளும் முன்னேற்றத்திற்கும் இத்தொழில் நுட்பத்தைப் பற்றி நன்கு புரிந்துகொள்ளவும் வழிவகுத்தன.

இயல் 5

இறுதியுரை

தமிழில் எழுத்துப் பிழை திருத்தியும் இலக்கணப் பிழைத் திருத்தியும் உருவாக்குவதற்கான வழிமுறைகளை ஆய்வதற்காக செய்யப்பட்ட ஆய்வு இது ஆகும். இத்தகைய திருத்திகளின் தேவை எப்போதோ உணரப்பட்டு இதற்காண முயற்சிகள் பல மேற்கொள்ளப்பட்டன. இத்தகைய முயற்சிகள் பல சரியான திசையில் செல்லாததன் காரணமாக கைவிடப்பட்டன அல்லது சரியான விளைவுகளை ஏற்படுத்தவில்லை. மொழியியல் அடிப்படையில் இத்தகைய ஆய்வுகள் அமையாதது இதற்கு முக்கியகாரணமாகும். எந்த ஆய்வும் அடிப்படையான தகவலகளை நன்கு தெரிந்த பின்னர்தான் மேற்கொள்ளப்படவேண்டும். எழுத்துப் பிழை மற்றும் இலக்கணப் பிழை திருத்தி உருவாக்க ஆய்வும் இதற்கு விதிவிலக்கல்ல. எவ்வாறு நோயை நன்கு புரிந்துகொண்டு சிகிச்சை ஆரம்பிக்கப் படுகின்றதோ அதுபோல் தான் மொழியில் நுட்பங்களை நன்கு புரிந்தபின் தான் இத்தகைய ஆய்வுகள் மேற்கொள்ளப்படவேண்டும். கணினியின் திறனை மட்டும் நம்பிக்கொண்டு செய்யப்படும் இத்தகைய ஆய்வுகள் குறை உடையதாகவும் பயனின்றியும் தான் அமையும். மொழியியல் மொழியின் இலக்கணத்தை நன்கு ஆய்ந்து தேவையான தகவலகளைத் தருகின்றது. அத்தகைய தகவல்கள் பிழை திருத்த ஆய்வுக்கு முக்கியமானதாகும். இத்தகைய ஆய்வு அடிப்படைகள் தான் திறமைமிக்க திருத்திகளை உருவாக்கத் துணை புரிய இயலும்.

தமிழில் பிழை திருத்திகள் பல உருவாக்கப்பட்டுள்ளன என்பதில் ஐயம் இல்லை. இவைகள் யாவும் எழுத்துப் பிழை திருத்திகள் என்ற நிலையில் தான் அமைகின்றன. இலக்கணப் பிழை திருத்தி என்று கூறத்தக்க கருவிகள் இல்லை என்று கூறலாம். எழுத்துப் பிழை திருத்திகளே இலக்கணப் பிழை திருத்திகளாகவும் உரிமை கோரப்படுகின்றன. இலக்கணத் திருத்தி உருவாக்கம் மிகப்பெரிய செலாகும். கணிப்பொறிக்கு தமிழ் இலக்கணத்தின் நெளிவு சுழிவுகளைக் கற்றுத் தந்தால்தான் இலக்கணத் திருத்தியாகக் கணினி செயலாற்ற இயலும். தரவுத்தொகுதிகளைக் கொண்டு கணிப்பொறி கற்றல் அணுகுமுறை அடிப்படையில் (machine learning approach) இத்தகைய செயல்பாட்டை எளிதாக இயலும். இதை தரவுத்தொகுதி இயக்க அணுகுமுறை

(corpus driven approach) என்றும் புள்ளியியல் அணுகு முறை (statistical approach) என்றும் கூறலாம்.

இரண்டுவிதமான அணுகுமுறைகள் உள்ளதாக முன்னர் பார்த்தோம்: (1) தரவுத்தொகுதி அடிப்படையிலான அணுகுமுறை (முற்றுநிலைத் தானியங்கியைப் பயன்படுத்தி) மற்றும் (2) மொழியியல் அடிப்படையிலான அணுகுமுறை (உருபனியல் ஆய்வியைப் பயன்படுத்தி). முதல் அணுகுமுறையில் தவறான சொல் தரவுத்தொகுதியின் மூலம் சரிபார்க்கப்பட்டு சரியான சொல் பரிந்துரைக்கப்படும். இரண்டாவது அணுகுமுறையில் தவரான உருபனியல் பகுப்பாய்வு சொல்லின் தவறை வெளிப்படுத்தி சரியான சொல்லைப் பரிந்துரை செய்யும். பரிந்துரையில் முக்கியமாக திருத்து தூரம் கணக்கில் கொள்ளப்படும். முன்னர் கூறியது போன்று பிழையானது.

பிழைகளை விடுபடு பிழை, செருகல் பிழை, பதிலீட்டுப் பிழை, இடமாற்றுப் பிழை என்று நான்காகப் பகுத்தாயலாம் என்று கண்டோம். இந்நான்கில் ஒரு பிழை மட்டும் கடுத்தில் கொள்ளப் பட்ட சொல்லில் இருந்தால் இதற்கான பரிந்துரை எளிதாக அமையும். அனால் ஒன்றிற்கும் மேற்பட்ட வகை பிழைகள் இருந்தால் பரிந்துரைகள் சிக்கலாக அமையும்.

பிழை ஆய்வை பின்பலமாக வைத்துக்கொண்டு பிழை திருத்தி உருவாக்குவது மிகுந்த பலனை அளிக்கும். இலக்கண விதிகளை வைத்துக்கொண்டு இவ்விதிகள் எவ்வாறு மீறப்படுகின்றன என்ற அடிப்படையில் இலக்கணத்திருத்திகள் உருவாக்குவது சரியான அணுகு முறையாகும். எழுத்துப் பிழை திருத்திகளைப் பொறுத்தவரியில் தமிழ் மொழியின் ஒலியனியல் அமைப்பை நன்றாக விளங்கிக்கொண்டு அதன் அடிப்படையில் எழுத்துப் பிழை திருத்தியை உருவாக்குவது நன்மை பயக்கும். இதில் சிக்கலான தகவல் என்னவென்றால் எல்லா விதிகளுக்கும் விதி விலக்குகள் காணப்படும். மொழி பிற மொழிகளிலிருந்து கடன் வாங்குவதன் காரணமாக அதன் தனித்தன்மையான ஒலியனியல் அமைப்பு சிதைவுற வாய்ப்புண்டு. தமிழில் சில எழத்துக்கள் மொழி முதலில் வரும், சில எழுத்துக்கள் மொழி இடையில் வரும், சில எழுத்துக்கள் மொழி இறுதியில் வரும் என்பன போன்ற ஒலியனையல் கட்டுப்பாடுகள் உண்டு. மேலும் மெய்மயக்கங்களின் வருகை

முறையும் விதிகளுக்குக் கட்டுண்டு வரும். இத்தகைய மொழி தொடர்பான தகவல்கள் திருத்திகள் உருவாக்கத்திற்குக் கைகொடுக்கும்.

துணைநூல் பட்டியல்

இந்திய மொழிகளின் நடுவண் நிறுவனம். 2004. தமிழ் நடைக் கையேடு. அடையாளம், மொழி டிரஸ்ட். சென்னை.

இராசேந்திரன், ச. **2002.** மொழி ஆய்வில் கணிப்பொறியின் பயன்பாடு. சித்திரபுத்திரன் மற்றும் பிறர் (பதிப்பு). ஆய்வியல் கட்டுரைகள். அன்னை அகரம் பதிப்பகம், தஞ்சாவூர்.

திருமுருகன், இரா. 2001. இனிய தமிழைப் பிழையின்றி எழுத எளிய வழிகள். பாவலர் பண்ணை, புதுச்சேரி.

நடராச பிள்ளை ந, விமலா ச. 1981. பிழை ஆய்வு: மொழி கற்பித்தலில் ஒரு புதிய பார்வை. தென்னிந்திய மொழிகளின் பயிற்று மையம் (இந்திய மொழிகளின் நடுவண் நிறுவனம்), மானசங்கோத்திரி, மைசூர்.

A Smart Spell checker system http://www.coe.neu.edu/

Anandan, ,Geetha, T.V. and Ranjani Parthasarathi. 2001.Morphological generator for Tamil. Tamil Inaiyam 2001. Malaysia.

Anandan, P, Ranjani Parthasarathy, Geetha, T.V. 2001. "Morphological Analyzer for Tamil." ICON 2002, RCILTS-Tamil, Anna University, India.

Attwell, E. and Elliott, S. 1987. Dealing with ill-formed English text, in: The Computational analysis of English. Longman, 1987.

Bellman, R. 1957. Dynamic Programming. Princeton University Press, Princeton, NJ.

Brill, E. 1992. "A Simple Rule-based Part of Speech Tagger," Proceedings of ANLP-92, 32rd Conference on Applied Natural Language Processing, Trento, Italy.

Damerau, F.J. 1964. "A technique for computer detection and correction of spelling errors." Communications of the ACM, 7 (3),171-176.

Dhanabalan, T. Ranjani Parthasarthi, Geetha, T.V. 2003. "Tamil Spell Checker", Tamil Inaiyam 2003, 18-27, Chennai, Tamilnadu.

Chaudhuri, B.B. 2002. "Reversed word dictionary and phonetically similar word grouping based spell-checker to Bangla text." Indian Statistical Institute, Kolkata, India.

Indian Language Spell Checker Design Workshop at DVPR Unit of Kolkata on 18-19th July, 2002.

Jensen, K, Heidorn,G.E., Richardson S.D. (eds.) 1993. Natural language Processing : The PLNLP approach, Kluwer Academic Publishers.

Jurafsky, D and Martin J.H. 2000. Speech and Language Processing. Pearson Education, Aisa.

Kernighan M.D., Church, K.W., and Gale, W.A. 1990. A Spelling correction program base on a noisy channel model. In COLING-90, Helsinki, Vol. II, pp. 205, 211.

Karlsson, Fred (ed.) 1995. Constraint Grammar: a language independent system for parsing unrestricted text. Mouton de Gruyter.

Lehmann, Thomas. 1993. A Grammar of Modern Tamil. second edition. Pondicherry Institute of Linguistics and Culture.

Levenshtein, V.I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. Cybernetics and Control Theory, 10 (8), 707-710. Original in Doklady Akademii Nauk SSSR 163 (4): 845-848 (1965).

Naber, D. 2003. A Rule-Based Style and Grammar Checker. Diploma Thesis, Computer Science – Applied, University of Bielefeld.

Rajendran, S. 2000. Spell and grammar checker for Tamil. UGC project report. Tamil University, Thanjavur.

Rajenran S, Arulmozi S, Ramesh Kumar S, & Viswanathan S. 2003. Computational Morphology of Verbal Complex. (Co-authors:) Language in India 3:4, www.languageinindia.com, April 2003, 376-398.

Rajendran, S 2006. “Parsing in Tamil: Present State of Art”. Indian Linguistics vol 67 2006, pages 159-167.

Rajendran S. 2006. “Shallow Parsing in Tamil: the state of Art”. Language in India, www.languageinindia.com.

Sengupta and Chaudhuri, B.B. 1996. "Morphological processing of Indian Languages for lexical interaction with application to spelling error correction". Sadhana,Vol.21, Part 3, pp 363-380.

Vaishnavi Ramaswamy. 2000. A morphological generator for Tamil. PhD Thesis. CALTS, University of Hyderabad, Hyderabad.

Vijay Sundar Ram. 2000. Developing a Tamil Spell Checker. AU-KBC Research Centre, MIT, Anna university.

Wagner, R.A. and Fischer, M.J. 1974. The string-to-string correction problem. Journal of the Association for Computing Machinery, 21, 168-173.

Walker , D.J., Clements, D.E., Darwin, M., and Amtrum, J.W. 2001. Sentence Boundary Detection: A Comparison of Paradigms for Improving MT Quallity, In Proceedings MT Summit VIII. Santiago de Compostela, Spain.